Vol. 48 NOVEMBER, 1957 No. 7

The Journal of Educational Psychology

UNIVERSITY OF MICHIGAN
JAN 30 1953
PERIODICAL READING ROOM

CONTENTS

Consistency of Sociometric Status Position in Male and Female High School Students 385
RITA R. WERTHEIMER
An Experimental Study of the Effect of External Direction During Learning on Transfer and Retention of Principles 391
JACK E. KITTELL
The Personality of Junior High and High School Students Attending a Residential School for the Blind 406
HERBERT M. GREENBERG, LOUISE ALLISON, MILDRED FEWELL, AND CHARLES RICH
A Procedure for Managing Teaching-Learning Situations in Educational Psychology 411
RAY H. SIMPSON
Sex Differences Among Teachers in the Use of the Criteria of Adjustment 426
HARRY BEILIN AND EMMY WERNER
Chronological Divisions of the Life Span 437
HORACE B. ENGLISH
A Comparison of Reading Comprehension Scores Obtained Before and After a Time Announcement 440
DESMOND L. COOK
Under- and Over-Age Grade-Placements and School Achievement 447
JACK A. HOLMES AND CARMEN J. FINLEY
Item Difficulties Based on End Segments 457
EDWARD L. CLARK
The Interrelationship of Area Test Scores and Cooperative General Culture Test Scores 460
JOHN C. WRIGHT AND BARRON B. SCARBOROUGH
Book Reviews 464
Publications Received 475

Published monthly except June to September
$7.00 a year in U. S. and Pan America; Canada, $7.20; other countries, $7.40.
Single issues $1.20

AMERICAN PSYCHOLOGICAL ASSOCIATION

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879.

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

WM. CLARK TROW, *Chairman*
University of Michigan

STEPHEN M. COREY
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

*Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

PAUL BLOMMERS
CHESTER W. HARRIS
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street, N. W.
Washington 6, D. C.

Copyright 1958 by the American Psychological Association, Inc.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume 48 November, 1957 Number 7

CONSISTENCY OF SOCIOMETRIC STATUS POSITION IN MALE AND FEMALE HIGH SCHOOL STUDENTS[1]

RITA R. WERTHEIMER

Western Psychiatric Institute, University of Pittsburgh

The present study represents an attempt to investigate the consistency of male and female sociometric status position through the use of sociometric retests spaced at intervals of 12 months, 20 months, and 8 months. A separate evaluation of male sociometric status consistency and female status consistency for these same retest intervals is also included. The sociometric criterion situations employed are in the general area of leisure-time activities. Product-moment correlation coefficients were calculated to evaluate the consistency of the total sociometric choice frequencies obtained in the several testings.

Previous studies have shown that sociometric test scores have a very high degree of consistency, particularly when the retest interval is short. Jennings (**2**), as part of an extensive study of sociometric status in a training school for delinquent girls, reports a coefficient of .96 on a retest of one housing unit at the school after a four-day interval. Newstetter, Feldstein and Newcomb (**3**) found an average coefficient of .95 based on data given on four successive weeks with a five choice allowance on one criterion (tent mates in summer camps). Zeleny (**5**) reported, at the college

[1] This study was conducted while the author was a member of the psychology department of West Virginia University. Grateful acknowledgment is made to Quin F. Curtis, Chairman of the psychology department at West Virginia, Delmas Miller, Principal of University High School, and to the many teachers who cooperated in this project.

level using five choices and one criterion (membership in a discussion group), coefficients ranging from .93 to .95 on readministering a sociometric test on successive days.

These research findings as a whole suggest that a high correlation coefficient obtained by the method of immediate retest is probably indicative of a memory factor, as well as the slowness with which people change in their choice behavior.

With a longer retest interval, coefficients go down as would be expected on an a priori basis. Bronfenbrenner (**1**) in a study with the entire population of the University of Michigan laboratory experimental school (with a range of students from nursery school to the sixth grade) found that rank correlation coefficients between raw status scores obtained by each student in the fall and spring sociometric testing varied from .27 for the nursery school group to .59 for the sixth grade group. He concluded that in classroom situations marked shifts in sociometric status were comparatively rare and that, on the whole, children tended to retain the same general social position, with this tendency more pronounced in the older age groups. Jennings (**2**, p. 235) spaced a sociometric retesting at an interval of eight months' distance from the original testing and found with her sample of delinquent girls a reliability coefficient of .72 for leisure-time companion choices.

Present Study

Sample Characteristics

All of the tenth- and eleventh-grade boys and girls from one consolidated high school were used in the original sociometric testing in March 1953. Retesting of these adolescents occurred in March 1954 and again in November 1954. Since sociometric choices in this study were limited to members of home room groups, the home room assignment policy adhered to in this high school is pertinent. Typically, students were given home room assignments in the tenth grade, and they retained the same home room grouping and teacher until they graduated from the school or were otherwise separated from the institution. Because of this the home room grouping presents a fairly stable social situation which should be well structured by the spring of the tenth grade, and a significantly high relationship between sociometric status position in each of

TABLE 1

SAMPLE CHARACTERISTICS AT THE TIME OF EACH SOCIOMETRIC TESTING: NUMBER OF CASES, SEX, MEAN AGE, AND GRADE LEVELS

Characteristics	March 1953 Sociometric Test	March 1954 Sociometric Test	November 1954 Sociometric Test
Total Number	200	200	97
Males	103	103	50
Females	97	97	47
Mean Ages (in yrs.)	16.3	17.3	18.5
Number of Grade Levels			
Tenth	97	0	0
Eleventh	103	97	0
Twelfth	0	103	97

the testing periods should exist. Table 1 summarizes some relevant data about the subjects.

Sociometric Measuring Instrument

The same near-sociometric questionaire previously used by Wertheimer (**4**) on a comparable high school group was employed. This near-sociometric questionnaire had been constructed in terms of the age and experience of high school adolescents. It was evident that the criteria on which choices were to be made had to be meaningful to them in terms of their own interests and experiences. The questions used were confined to roughly comparable social situations. The attempt was also made to keep the questions on the same "level" of social rapport. For example, it was conceivable that an acceptable companion at a school party might not be the same one chosen for discussion of a serious personal problem Additionally, all of the questions were phrased positively so that the adolescents would not have their attention called to the less favored members of each group and, in this way, cause even more severe social maladjustment to these students.

The sociometric questionnaire that was used included the following instructions to the students:

> What we are going to ask you to do today is not a test of any kind, and will not become a part of your school record. There are no right or wrong answers to it. Anything that you include here will be confidential and will only be used as research material.

In response to the following questions, be sure and write down the three persons' first and last names. You may put down the same name in response to more than one question, if you would really choose the same person.

No additional specific instructions were given to the students. They were able to make companion choices from either sex in all three sociometric situations. The three criterion situations for which three choices were to be given were:

1. Suppose you have just won first prize on a quiz program. Among the prizes is an expense-free trip to Canada for yourself and three students *in your home room group*. Whom would you choose to go along with you?
   a. _____ b. _____
   c. _____
2. Suppose you are planning a party at your home and have to limit the number of people you can ask to it. Which three members *of your home room class* would you invite?
   a. _____ b. _____
   c. _____
3. Suppose you have a chance to spend the summer at a camp which has a section for girls and a section for boys. Which three members *of your home room group* would you prefer as fellow campers?
   a. _____ b. _____
   c. _____

Administration of Questionnaire

Since all of the home room classes in the tenth and eleventh grades of the high school (in the March 1953 testing) were requested to fill in the questionnaire, it was necessary that its administration be closely geared to the school schedule. In consultation with the school authorities, it was decided that it could be most expeditiously administered by the teachers themselves. Accordingly, the experimenter met with them as a group, commented on the questionnaire procedure, possible sources of difficulty, etc. They were given the questionnaires in large envelopes that they were requested to seal in the presence of the students after administration and to return the material to the school office on the same day. With the exception of specific instructions to the students to limit their choices to their own home room groups, the teachers reported no difficulties in connection with the administration procedure.

The same procedure and sociometric questionnaire were used in the March 1954 and November 1954 administrations.

Sociometric Questionnaire Analysis

A roster of each home room class was obtained and a frequency tabulation made of the number of times each home room member was chosen by any other member as a companion for any or all of the criterion situations. Each person in the home room group obtained a total sociometric choice frequency based on three choices for each of the three leisure-time activity situations suggested in the questionnaire.

Treatment of the Data and Results

To evaluate the consistency of these total sociometric choice frequencies obtained by the designated students (both as a group and in terms of the male and female dichotomy) on subsequent sociometric testings, Pearson product-moment correlation coefficients were calculated between (*a*) March 1953 and March 1954 data for the total group and for the males and females separately; (*b*) March 1953 and November 1954 data for the entire group and for males and females separately; and (*c*) March 1954 and November 1954 data for the total group and for males and females separately.

The obtained correlation coefficients are presented in Table 2.

All of the nine coefficients are statistically significant beyond the .01 level and vary in size from .55 to .71. These data suggest that there is a high degree of consistency in sociometric status positon (as shown in expressed preferences for leisure-time com-

TABLE 2

PRODUCT MOMENT CORRELATIONS OF TOTAL SOCIOMETRIC FREQUENCIES OF MALE AND FEMALE HIGH SCHOOL STUDENTS IN THREE SOCIOMETRIC TESTINGS

Groups	March 1953–March 1954 Comparisons			March 1953–November 1954 Comparisons			March 1954–November 1954 Comparisons		
	N	*r*	*df*	*N*	*r*	*df*	*N*	*r*	*df*
Total	200	.55	198	97	.66	95	97	.68	95
Male	50	.56	48	50	.69	48	50	.67	48
Female	47	.62	45	47	.62	45	47	.72	45

Note.—All of these correlation coefficients are significant beyond the .01 level for the degrees of freedom indicated in each case.

panions) in adolescent high school students as a group and that this consistency is maintained by both males and females.

While this consistency in choice position may vary somewhat in different high schools (as, for example, in home room grouping systems that are more fluid than the system in the present study), it is likely that social status in leisure-time activities is well structured in the tenth grade and is probably accentuated but at least maintained until the last year in high school. These data further suggest that a pronounced group development had occurred by the tenth grade in the sample used in this study.

Summary

The present study investigated the consistency of male and female adolescent sociometric status position through the use of sociometric retests spaced at intervals of 12 months, 20 months, and 8 months. A separate evaluation of male sociometric status consistency and female sociometric status consistency for these same retest intervals was also included. The sociometric criterion situations employed were in the general area of leisure-time activities. Product-moment correlation coefficients were calculated to evaluate the consistency of the total sociometric choice frequencies obtained on the several testings, and all of these coefficients were statistically significant beyond the .01 level. The obtained results suggest that there is a high degree of consistency in sociometric status position (as shown in expressed preferences for leisure-time companions) in adolescent high school students as a group and that this consistency is maintained by both males and females.

References

1. Bronfenbrenner, U. The Measurement of sociometric status, structure, and development. *Sociometry Monogr.*, 1945, **6**, 3–80.
2. Jennings, H. H. *Leadership and isolation.* New York: Longmans, 1950.
3. Newstetter, W. I., Feldstein, M. J., & Newcomb, T. M. *Group adjustment.* Cleveland: Western Reserve Univer. Press, 1938.
4. Wertheimer, R. R. The relationship between sociometric status, sex, and socioeconomic level and Rorschach signs of adjustment with a group of adolescent subjects. Unpublished doctoral dissertation, Univer. of Pittsburgh, 1953.
5. Zeleny, L. D. Sociometry of Morale, *Sociol. Rev.* 1939, **6**, 799–808.

Received December 2, 1956.

AN EXPERIMENTAL STUDY OF THE EFFECT OF EXTERNAL DIRECTION DURING LEARNING ON TRANSFER AND RETENTION OF PRINCIPLES

JACK E. KITTELL

University of California, Berkeley

It is not unusual for two extremes to be associated with educational practice: (*a*) those teaching-learning situations containing such a minimum of structuring that they can be characterized as relatively independent discovery situations, and (*b*) those situations containing complete presentations of selected, specific principles and facts and characterized as rote memory situations. Various research findings have indicated that, perhaps, amounts of direction during learning that lie between those two extremes may be of value in promoting maximum learning, transfer, and retention when used in classroom situations. The results of this investigation support suggestions that some direction during learning aids discovery, transfer, and retention of principles.

The Problem

Experimental research has produced evidence that learner activity is of value in the discovery of principles. Classroom provisions for active pupil discovery have been found to be superior to methods emphasizing the specification of correct responses by the teacher in advance of discovery (**1, 10, 13, 14**). However, the extreme proposition that the more the freedom of learner activity and the less the amount of external direction, the better, as has been inferred in educational literature (**11,** p. 141; **15,** p. 80) is not necessarily a valid one. Various experiments have demonstrated that some degree of direction in discovery is superior to none at all in promoting transfer (**7, 8, 9**).

Three recent experiments somewhat similar in design and purpose to the one reported here have produced conflicting inferences. One experiment, using sixth-grade pupils as *S*s, produced results favoring independent discovery as superior to specification of correct response in advance of discovery activities (**12**).

Two other recent experiments used college graduates and college students as *S*s. The first of these two experiments found that of four different degrees of direction during discovery, ranging from a control group receiving no direction to a group supplied underlying principles but no correct responses, the latter was superior in promoting transfer (**4**). The other experiment compared the effects of supplying principles as direction and independent discovery on transfer and retention (**3**). No significant differences were found in the discovery of new principles. After an interval of one month the group supplied with principles during discovery retained significantly more learned principles than the control group.

The results of these experiments suggest that direction, in the form of underlying principles, supplied during discovery is beneficial in promoting transfer to similar problems and to retention after one month. It would seem important to conduct further experiments with elementary school pupils in situations somewhat comparable to actual teaching-learning situations to obtain evidence of the relative effects of Minimum, Intermediate, and Maximum amounts of direction during discovery on transfer to situations of varying degrees of differentness and on retention.

The hypotheses tested in the present experiment were that the amount of direction characterized as Intermediate (*a*) effects greater increases in learning and transfer to three situations of varying differentness, and (*b*) promotes greater retention through periods of two and four weeks.

Materials and Procedures

Subjects

The *S*s of this experiment, 132 sixth-grade pupils from public elementary schools in Pullman, Washington, were placed into three treatment groups through stratified-random selection. Stratification was based on high, medium, and low reading achievement classifications.

Final membership in the *Minimum* treatment group numbered 45, in the *Intermediate* treatment group, 44, and in the *Maximum* treatment group, 43.

Treatment group means and standard deviations of *S.R.A. Primary Mental Abilities* intelligence quotients and chronological ages are shown in Table 1.

TABLE 1

SRA INTELLIGENCE QUOTIENT AND CHRONOLOGICAL AGE MEANS AND STANDARD DEVIATIONS FOR THREE TREATMENT GROUPS

Statistic	Treatment Groups		
	Minimum	Intermediate	Maximum
Number of cases	45	44	43
IQ mean	107.0	105.98	106.30
Standard deviation	13.09	12.55	12.84
C. A. mean (months)	131.26	131.68	132.33
Standard deviation	4.55	4.48	5.06

TABLE 2

MEANS AND STANDARD DEVIATIONS OF NUMBER OF PRINCIPLES KNOWN ON PRETRAINING TEST FOR THREE TREATMENT GROUPS

Statistic	Treatment Groups		
	Minimum	Intermediate	Maximum
Mean	1.444	1.864	1.630
Standard deviation	1.500	1.540	1.100

The statistical difference among treatment-group IQ means were tested by referring the ratio of the variances among group means to the variance within groups to a table of the *F* distribution. The *F* of .063 was obtained indicating no significant differences among group means.

The distribution of chronological ages could not be accepted as normal and a nonparametric test, analysis of variance by ranks, was applied to test the significance of differences among the three treatment groups. The *H* value, 1.03, was found to be not significant. *H*, with the group numbers used in this experiment, approximates a chi-square distribution.

The mean number of principles known on the pretraining test and standard deviations for the three treatment groups are presented in Table 2.

The distribution of number of principles known on the pretraining test could not be accepted as normal, and again analysis of variance by ranks was applied to test the significance of differences among the three treatment groups. A resulting *H* of .0526 was found not to be significant.

On the basis of the above analyses it was assumed that the three treatment groups did not differ significantly in initial mental ability or pretraining knowledge of the principles used in the experiment.

Pretraining and Posttraining Tests

The pretraining and first posttraining tests for knowledge of principles consisted of 45 items grouped by threes for each of 15 organizational principles.

A second posttraining test, consisting of 45 items different from any used on other tests and grouped by threes for each of the same 15 principles, was used to measure transfer to new items based on previously experienced principles.

A third posttraining test, consisting of 45 different items grouped by threes for each of 15 different organizational principles, was used to measure posttraining ability to discover and use unlearned principles.

An example of the type of verbal problems used throughout this experiment follows: (*a*) Gone, (*b*) Start, (*c*) Go, (*d*) Stop, (*e*) Come.

The five words of each item were selected so that four were related by some principle and one did not belong. The principle was the reason why one word did not belong and it became the learner's task to discover the principle and apply it in order to discover the word not belonging. In the example the principle was that of paired opposites. "Start" and "Stop" are direct opposites as are "Come" and "Go" but "Gone" is of the wrong tense to be a direct opposite of any of the other four and so does not belong.

Twelve principles used in the pretraining, the training, and the posttraining tests were based on meaning relationships, one was based on sound relationships, one was based on familiar word combinations, and one was based on word arrangement and meaning relationships. If a wrong response was made for any item of a group of three, the *S* was given no credit for knowledge of the applicable principle.

The tests were distributed as booklets for group administration with all necessary instructions on the cover page. The directions stated that the items were grouped and spaced on the basis of underlying principles but gave no clue as to the nature of the

principles. The *Ss* were required to mark an X in the answer space corresponding to the word that did not belong with the other four in each item.

Seventy-six of the 180 items used in this experiment and 10 of the 30 principles had been previously used in a recent, somewhat similar experiment (**12**). All of the items used were composed of words selected from the first thousand of each of two well-known reading word lists (**5, 6**). Such selection provided reasonable assurance that the items were not tests of reading ability for the sixth grade pupils serving as *Ss*.

All three tests used for measurement in this experiment were administered to 87 sixth-grade pupils in another community prior to the initiation of this experiment. An alternate form procedure was used to determine reliability and after correcting for length of one test, a reliability coefficient of .74 was obtained.

The third posttraining test composed of different items based on 15 different principles was compared with the first posttraining test for relative difficulty. The percentage of correct responses for each item was computed revealing a range from 6% to 72% with a mean of 35% for the first posttraining test and from 6% to 79% with a mean of 23% for the third posttraining test. The same procedure was used on the basis of principles with a resulting mean of 8% for the pretraining test and a mean of 2.3% for the third posttraining test. Five educators acting as judges supported the accepted assumption that the third posttraining test was more difficult than the first posttraining test.

Learning Materials and Treatments

Items used in the training tests were different but of the same type as those used for pre- and posttraining tests. The principles used for training were the same as those used in the pre-, first, and second posttraining tests.

Three different combinations of clues to the principles determining correct responses to the training items were incorporated into three sets of training materials, which when administered, each to a different group of *Ss*, effected different treatments by providing three different amounts of direction during learning. The procedure of administration was consistent throughout the training period and followed practices previously tested by research. (**3, 4**).

The learning material for the group receiving Minimum direction consisted of items arranged as in the pre- and posttraining tests. Three items based on the same principle were grouped together and separated from other groups of items by spacing, The *S*s of the Minimum group were told that each group of three items was based on an underlying principle. No information was given this group as to the nature of the principles except that because of them one word in each item did not belong.

The group receiving Intermediate direction was provided with all the information supplied the Minimum group plus a verbal statement of the principle involved. The principles were printed immediately preceding the group of items to which they applied and were read orally twice with each administration of the items based on the principle. The stated principle for the item previously used as an example is: Form Two Pairs of Opposites.

The materials used by the Maximum treatment group included all the clues provided the Intermediate group plus oral statements of the three correct responses for each group of items. The oral statements of correct responses were made for one group of items at a time prior to the marking of responses by the *S*s. The oral statement made for the item previously used as an example was: "'Start' is the opposite of 'Stop' and 'Come' is the opposite of 'Go' so 'Gone' is the word that does not belong."

Knowledge of correctness or incorrectness of responses was provided all *S*s by the use of corrugated cardboard answer sheet backings, one surface layer of which was punctured to provide holes into which a pencil point dropped when the correct response for each item was marked. Such devices have been used successfully with groups in other experiments (**3, 4**). Their use in the present experiment was accompanied by adequate proctoring and instructions which included a statement that *S*s should continue to mark answer spaces until the correct response was made, but that each incorrect response mark would lower the *S*'s total score and so no unnecessary marks should be made.

The training period was five weeks in length with nine items based on three principles being presented and repeated twice each week for a total presentation of 45 items based on 15 principles. To discourage *S*s' attempts to rely on memory of correct response patterns for the items presented a second and third time each week, the words in each item were varied in order for each

presentation on the basis of five different correct-response patterns. The same patterns were used in punching holes in the answer sheet backings. The five different patterns were used in rotating order.

The pretraining test when administered the second time following training became the first posttraining test. The number of principles actually learned and transferred was the number of principles known on the first posttraining test not known on the pretraining test. Treatment-group means and variances of the number of principles learned and transferred to the first posttraining test are presented in Table 3.

Analysis of variance was used to test the significance of differences among the three treatment groups. The F, 4.73, was well within the .05 level of confidence. Accepting the hypothesis that significant differences existed among the three treatment groups, pairs of groups were compared through the use of a t test. Table 4 presents the differences between treatment-group means and the significance of t for each.

TABLE 3

TREATMENT-GROUP MEANS AND VARIANCES FOR NUMBERS OF PRINCIPLES LEARNED AND TRANSFERRED ON FIRST POSTTRAINING TEST

Statistic	Treatment Groups		
	Minimum	Intermediate	Maximum
Mean	2.555	3.660	3.744
Variance	4.290	3.570	4.440

TABLE 4

SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS OF TREATMENT GROUPS FOR PRINCIPLES LEARNED AND TRANSFERRED ON FIRST POSTTRAINING TEST

Group Pair	Difference Between Means	Standard Error of Difference	t	Level of Significance
Minimum and Intermediate	1.106	.422	2.620	P .02
Minimum and Maximum	1.189	.448	2.654	P .02
Intermediate and Maximum	.084	.429	.196	Not significant

Both the Intermediate and Maximum treatment groups were superior to the Minimum treatment group in learning and transferring principles to previously encountered situations. There was no significant difference between Intermediate and Maximum treatment groups in this situation.

Amount of Transfer to New Item Situations

The number of principles learned and transferred to new item situations was the number of principles known on the second posttraining test not known on the pretraining test. The mean number of principles learned and transferred and variances for each treatment group as measured by the second posttraining test are presented in Table 5.

An F of 18.87 was obtained through analysis of variance and found significant at the .001 level of confidence. The hypothesis that significant differences existed among the three treatment groups in the number of principles transferred to new item situations was accepted. The t test was applied to test significance of differences between pairs of treatment-group means. Table 6

TABLE 5

TREATMENT GROUP MEANS AND VARIANCES FOR NUMBERS OF PRINCIPLES LEARNED AND TRANSFERRED TO NEW ITEM SITUATIONS IN SECOND POSTTRAINING TEST

Statistic	Treatment Groups		
	Minimum	Intermediate	Maximum
Mean	1.930	4.590	2.907
Variance	3.520	4.160	5.190

TABLE 6

SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS OF TREATMENT GROUPS FOR PRINCIPLES LEARNED AND TRANSFERRED TO SECOND POSTTRAINING TEST

Group Pair	Difference Between Means	Standard Error of Difference	t	Level of Significance
Minimum and Intermediate	2.657	.415	6.40	P .001
Minimum and Maximum	.974	.444	2.19	P .05
Maximum and Intermediate	1.683	.458	3.68	P .001

presents differences between treatment-group means and the significance of t for each.

Inspection of Tables 5 and 6 reveals that the Intermediate treatment group was significantly superior to the other two treatment groups in the number of learned principles transferred to new item situations. The Maximum treatment group was superior to the Minimum treatment group in transferring principles to new item situations.

Posttraining Discovery of New Principles

The third posttraining test was composed of 45 items based on 15 principles. Neither items nor principles had been previously encountered by the *S*s during this experiment. The number of principles known on this test was determined by counting the number of principles for which *S* made three correct item responses. Treatment-group means and variances of the number of principles known on the third posttraining test are presented in Table 7.

The distribution of numbers of principles known on this test was considered questionable when the Chi-square Goodness of Fit test yielded a chi-square value of 15.03 found to be significant at the .02 level of confidence. Both a parametric technique, analysis of variance, and a nonparametric technique, analysis of variance by ranks, were applied to test the significance of differences among the three treatment groups as to the number of new principles discovered. An F of 38 and an H of 30.88 were obtained, both significant at the .001 level of confidence.

The t test and analysis of variance by ranks were both applied in testing for significance of differences between pairs of treatment groups. As both techniques produced values at the same levels of significance only the results of the t test are presented here.

TABLE 7

TREATMENT-GROUP MEANS AND VARIANCES FOR NUMBERS OF PRINCIPLES KNOWN ON THIRD POSTTRAINING TEST

Statistic	Treatment Groups		
	Minimum	Intermediate	Maximum
Mean	1.733	4.590	2.674
Variance	4.20	4.67	4.93

TABLE 8

SIGNIFICANCES OF DIFFERENCES BETWEEN TREATMENT GROUP MEANS OF NUMBERS OF PRINCIPLES KNOWN ON THIRD POSTTRAINING TEST

Group Pair	Difference Between Means	Standard Error of Difference	t	Level of Significance
Minimum and Intermediate	2.857	.446	6.41	P.001
Minimum and Maximum	.941	.4525	2.08	P.05
Maximum and Intermediate	1.916	.468	4.09	P.001

Table 8 presents the differences between treatment-group means of the number of principles known on the third posttraining test and the significance of t for each.

The Intermediate treatment group was able to discover a significantly greater number of new principles than either of the other two treatment groups. The Maximum treatment group was able to discover a significantly greater number of new principles than the Minimum treatment group.

Number of Principles Retained

The pretraining test administered for a second time at the end of the training period as the first posttraining test was administered a third time two weeks following its second administration and two weeks later it was administered for a fourth time to measure the number of principles retained for two and four weeks. The number of principles retained was determined by the number known on the first posttraining test and also known on the two- and four-week retention tests.

The numbers of principles retained by each treatment group were compared establishing that the Intermediate group retained significantly more principles than the other two treatment groups for two and four weeks and the Maximum group significantly more than the Minimum group. However, more indicative of the relative effects of the amounts of direction on retention were the comparisons of the retention loss by each treatment group.

Analysis of variance by ranks for matched samples was used to test the significance of retention losses. Table 9 presents the treatment-group means of numbers of principles known on the

TABLE 9

SIGNIFICANCE OF DIFFERENCES FOR EACH TREATMENT GROUP BETWEEN THE NUMBER OF PRINCIPLES KNOWN ON THE FIRST POSTTRAINING TEST AND THE NUMBER REPEATED AFTER TWO WEEKS

Treatment Group	Mean Number of Principles		χ_r^2	Level of Significance
	Originally Known	Repeated		
Minimum	2.555	1.60	15.22	P.001
Intermediate	3.660	3.50	.82	P.40
Maximum	3.744	2.86	20.03	P.001

Note.—Analysis of variance by ranks, matched samples.

TABLE 10

SIGNIFICANCE OF DIFFERENCES BETWEEN THE NUMBER OF PRINCIPLES RETAINED TWO AND FOUR WEEKS FOR EACH TREATMENT GROUP

Treatment Group	Mean Number of Principles Retained		χ_r^2	Level of Significance
	Two Weeks	Four Weeks		
Minimum	1.60	1.177	5.60	P.02
Intermediate	3.50	3.295	1.15	P.30
Maximum	2.86	2.300	11.12	P.001

Note.—Analysis of variance by ranks, matched samples.

first posttraining test, those retained after two weeks, χ_r^2 values, and the level of significance for each.

The Intermediate treatment group made no significant loss in retention of principles the first two weeks. Both the Maximum and Minimum treatment groups made significant losses in numbers of principles retained for two weeks.

Table 10 presents the treatment-group means of numbers of principles known on the two-week retention tests and repeated on the four-week retention test, the χ_r^2 values and the level of significance for each.

Both the Maximum and Minimum treatment groups continued to make significant losses in retention of principles while the Intermediate treatment group still had made no significant loss in the number of principles retained.

The significance of the losses of retention of principles by each

TABLE 11

SIGNIFICANCE OF DIFFERENCES FOR EACH TREATMENT GROUP BETWEEN NUMBERS OF PRINCIPLES KNOWN ON FIRST POSTTRAINING TEST AND THE NUMBERS REPEATED AFTER FOUR WEEKS

Treatment Group	Mean Number of Principles		χ_r^2	Level of Significance
	Originally Known	Repeated		
Minimum	2.555	1.177	20.00	P.001
Intermediate	3.660	3.295	2.75	P.10
Maximum	3.744	2.320	25.84	P.001

Note.—Analysis of variance by ranks, matched samples.

treatment group for the total four-week retention period is presented in Table 11.

The Intermediate treatment group made no loss in retention of principles during the four-week retention period that was significant at the .05 level of confidence. Both the other two treatment groups made significant losses in the retention of principles during both of the two-week retention periods and for the total of four weeks.

Discussion

In the present experiment, sixth-grade pupils benefited from clues which, in addition to organization of materials and information that such organization existed, provided relationships on which to base discovery.

Reasonable interpretation of all the evidence available indicates that *S*s benefit from help given them in their search for bases determining correct responses. Specification of answers in advance, however, seems to encourage reliance on rote memory rather than discovering and applying on the basis of underlying relationships. Evidence provided by the present investigation seems to support such indications.

The present investigation differed from previous experimentation in that the relative effects of Minimum, Intermediate, and Maximum amounts of direction on transfer to three different situations and on retention were investigated within a single design. The Intermediate amount of direction, which provided as clues to discovery statements of underlying principles, proved to be superior

to a Minimum amount of direction in promoting learning and transfer to all three types of situations. The Intermediate amount of direction was superior to Maximum amount of direction in promoting transfer to all but one type of situation. Two and four weeks after training, the group receiving Intermediate direction retained significantly more principles known in that situation than did the groups receiving Maximum or Minimum amounts of direction. The evidence provided by this experiment indicates that both the Minimum and Maximum amounts of directions during learning became less and less effective as the situations to which transfer was made became increasingly different and as the elapsed time between training and measurement of transfer became greater. The Intermediate amount of direction increased in effectiveness between the first and second transfer situations and maintained the new level of effectiveness on the third transfer situation.

The scope of this experiment has acknowledged limitations. Should further research confirm the following implications, the teacher's role in providing direction in learning-teaching situations would be clarified.

Evidence provided by the present experiment suggests that:

1. In addition to organizing the materials used in learning, teachers should aid pupil discovery by suggesting meaningful relationships on which learners may base discovery and by providing practice with those relationships.

2. Providing statements of underlying relationships without specifying answers fosters learning, retention, and transfer to different situations.

Summary

The purpose of this experiment was to determine the relative effects of three amounts of direction to learners in their discovery of established principles on transfer to differing situations and on retention of learned principles.

Each of three groups of sixth-grade pupils was supplied a different amount of direction during discovery of principles determining solution of multiple-choice verbal items. The group receiving an Intermediate amount of direction consisting of organized materials, information of such organization, and statements of underlying

relationships, learned and transferred as many or more principles to three differing situations than the groups receiving less or more direction. After periods of two and four weeks subsequent to training, the group receiving an Intermediate amount of direction retained a greater proportion of learned principles than the other two groups.

Evidence from this experiment in conjunction with that of similar experiments indicates that furnishing learners with information in the form of underlying principles promotes transfer and retention of learned principles and may provide the background enabling future discovery of new principles.

References

1. Anderson, G. L. Quantitative thinking as developed under connectionist and field theories of learning. In Esther G. Swenson, G. L. Anderson, & C. Stacey (Eds.), *Learning theory in school situations*. Minneapolis: Univer. of Minnesota Press, 1949. Pp. 40–73.
2. Brownell, W. A., & Moser, H. E. *Meaningful versus mechanical learning: a study in grade III subtraction*. Durham, N. C.: Duke Univer. Press, 1931.
3. Craig, R. C. Directed versus independent discovery of established relations., *J. educ. Psychol.*, 1956, **47**, 223–234.
4. Craig, R. C. *The transfer value of guided learning*. New York: Bureau of Publ., Teachers Coll., Columbia Univer., 1953.
5. Dolch, E. W. *Problems in reading*. Champaign, Ill.: Garrard Press, 1948.
6. Gates, A. I. *Reading vocabulary for primary grades*. (rev. ed.) Bureau of Publ., Teachers Coll., Columbia Univer., 1935.
7. Hendrickson, G., & Schroeder, W. H. Transfer of training in learning to hit a submerged target. *J. educ. Psychol.*, 1941, **32**, 205–213.
8. Judd, C. H. The relation of special training to general intelligence. *Educ. Rev.* **36**, 1908, 28–42.
9. Katona, G. *Organizing and memorizing*. New York: Columbia Univer. Press, 1940.
10. McConnell, T. R. Discovery versus authoritative identification. In T. R. McConnell et al. (Eds.), *Studies in the psychology of learning II*. Iowa City: Univer. of Iowa, 1934.
11. Spitzer, H. F. Learning and teaching arithmetic. In N. B. Henry (Ed.), *Yearb. nat. Soc. Stud. Educ.*, **50**, Part II. *The Teaching of Arithmetic*. Chicago: Univer. of Chicago Press, 1951.
12. Stacey, C. L. The law of effect in the retained situation with meaningful material. Unpublished doctoral dissertation, Univer. of Minnesota, 1945.

13. Swenson, Esther J. Organization and generalization as factors in learning, transfer, and retroactive inhibition. In E. J. Swenson, G. L. Anderson, & C. L. Stacey (Eds.), *Learning theory in school situations*. Minneapolis: Univer. of Minnesota Press, 1949. Pp. 9–39.
14. Thiele, C. L. *The contribution of generlization to the learning of the addition facts*. New York: Bureau of Publ., Teachers Coll., Columbia Univer., 1938.
15. Thiele, C. L. Arithmetic in the middle grades. In N. B. Henry (Ed.), *Yearb. nat. Soc. Stud. Educ.*, **50**, Part II. *The Teaching of Arithmetic*, Chicago: Univer. of Chicago Press, 1951. Pp. 76–102.

Received January 11, 1957.

THE PERSONALITY OF JUNIOR HIGH AND HIGH SCHOOL STUDENTS ATTENDING A RESIDENTIAL SCHOOL FOR THE BLIND

HERBERT M. GREENBERG, LOUISE ALLISON,

MILDRED FEWELL, AND CHARLES RICH

Texas Technological College, Lubbock

The method by which the blind can best and most profitably be educated has been a subject for controversy for many years. The impact on the personality of the blind child of the residential school which educates better than three-quarters of all blind children is widely debated in the field. Some work has been done regarding the intellectual effects (**5**), which in general has shown satisfactory intellectual growth for blind students. Nothing, however, has been done in the way of focusing on the school for the blind as an effecter of personality. It is the purpose of this study, through use of a sample from a representative school for the blind, to compare the norms of this group on several personality traits with the general norms of their sighted age mates. In other words, it is the purpose of this study to investigate a hypothesis that students attending a residential school for the blind will be: more neurotic, less self-sufficient, more submissive, less self-confident, about as sociable, and more authoritarian than the norm for the sighted group for these traits.

Procedure

Subjects

One hundred and three students, constituting practically all those attending grades six through twelve in a residential school for the blind, were tested. This school has two categories of students. By far the largest number live at the school for the approximate 9½-month academic year and return home for the summer. A few whose families live in the town in which the school is located return home daily or on weekends. Living at home is greatly discouraged by the school administration.

The students represent a cross section of backgrounds and a

reasonably accurate cross section of the socioeconomic groupings of the state. There is a wide range of academic ability, with the norm probably being very similar to that of the sighted population. Those few graduates who attend college are generally quite successful, though it should be noted that this number is quite small.

Test Battery

The Bernreuter Personality Inventory, the California F Scale, and a personal data form were administered to the students by administrators working simultaneously with four different groups.

The Bernreuter Personality Inventory has been standardized on thousands of high school and college students and adults. Percentile norms have been established for each of these groups, and for males and females in each group. The traits measured by this instrument are: neurotic tendencies, (B1N); self sufficiency, (B2S); introversion-extroversion, (B3I); dominance submission, (B4D); self confidence, (F1C); social ability, (F2S). The B3I was not employed in the present study, due to its very high correlation with B1N and F1C. The Bernreuter consists of 125 items to which the respondent answers with a yes, no, or question mark.

The California F Scale was constructed by Adorno and others, and has been standardized on several thousand students and adults. It consists of 29 statements to which the *S* responds by various shadings of agreement or disagreement (strongly agree +3, agree +2, mildly agree +1, strongly disagree −3, disagree −2, mildly disagree −1). An example of the type of items employed is "familiarity breeds contempt."

A personal data form has been employed by the senior author in his doctoral dissertation (**4**). It was designed to take data pertaining to the background of the subject so that the establishment of comparison groups would be possible. It contains such items as age, sex, approximate yearly income of family, religious persuasion, length of time at school, satisfaction with school, grades, etc. It is not in any way to be deemed a psychological test of a standardized instrument.

Methods of Collecting Data

In order to eliminate so far as possible the administration as a variable, a standardized administration was used by all four.

A script was prepared, which was followed exactly by each, and thorough rules as to the type of questions that could be answered were made and followed. Only definitions of words could be given to the *S*s and no interpretation of test items was permitted. Those *S*s who were able responded on typewriters to orally dictated questions while others employed pencil or Braille.

The respondents were told not to put their names anywhere on the paper and were assured that their responses could not be traced in any way to them. It was felt that with this guarantee of anonymity more candor might be hoped for.

Methods of Treating Data

Means and standard deviations were found for each of the three subgroups: Subgroup A—6th, 7th, and 8th grades; Subgroup B 9th and 10th grades; and Subgroup C 11th and 12th grades. These groups were further broken down into male and female for the Bernreuter B scales due to the differential sex standardization of these scales on the inventory. The t ratios were then taken between means of each subgroup for each of the traits tested. Results for each question of the personal data form were tabulated for each of the subgroups.

Results

Table 1 gives the mean percentile for the several subgroups for each of the traits measured. From this table it can be seen that the group exhibited high neurotic tendencies. The lowest subgroup had a mean score at the 66th percentile, while the remainder were at the 80th and above. It can also be seen that the group is quite dependent. One subgroup has a mean as low as the 15th percentile, while the most self-sufficient subgroup had its mean at the 34th percentile.

It can further be seen that this trend is carried out in the dominance-submissive scale, the groups lying on the submissive side. This was the only trait besides the authoritarian in which significant t ratios were found between any of the subgroups. In this case it was between the youngest and the oldest males, t ratios being 3.15, or above the 1% level of significance.

From the high neurotic trend exhibited by the group, it might be assumed that there is a lack of self-confidence. This, in fact, is clearly indicated by Table 1 which shows that no subgroup

TABLE 1

RESULTS FOR THE SEVERAL TESTED TRAITS BERNREUTER SCALES EXPRESSED IN PERCENTILE FIGURES[a]

Group	B1-N	B2-S	B4-D	F1-C	F2-S	F scale
Male A	82	15	15	81	29	152
Female A	66	22	41	81	29	152
Male B	83	19	23	86	25	167
Female B	83	17	37	86	25	167
Male C	81	21	46	80	45	151
Female C	80	34	54	80	45	151

[a] The first five traits on the above table are scales of the Bernreuter Personality Inventory taken from the manual of that inventory. The sixth trait refers to the California F scale.

B1-N—Neurotic Tendencies; B2-S—Self Sufficiency; B4-D—Dominance-Submission; F1-C—Self Confidence; F2-S—Social Ability; F scale—Authoritarianism.

has a mean below the 80th percentile, thus demonstrating a severe lack of self-confidence.

It should be noted that the trait of sociability is the only one in which the group lies on the healthy side of the norm. The most unsocial subgroup having a mean percentile of 45, however, authoritarianism is carried out in the pattern set by most traits, the means ranging from 151 to 167 on a scale ranging from 29 to 203, the high numbers representing high authoritarianism. It should be further noted that the 9th and 10th graders show a significantly more authoritarian score than the 6th, 7th, and 8th graders, the *t* ratios being 4.71 or again well above the 1% level of significance.

The most noticeable result obtained by the personal data form was the apparent widespread dissatisfaction with the school. In Subgroup A, 27 would choose this school again if having to choose again, while 17 would not; in Subgroup B, 13 would choose again while 17 would not; and in Subgroup C, 16 would choose again and 17 would not. Also, many who answered affirmatively did so because they felt that there was no alternative but expressed great hostility toward the school. The subgroups answered the questions quite consistently with one another. There were no differences in the responses to the other questions on the personal data form as great as the difference indicated in the above para

graph except, of course, for the obvious differences of age and length of attendance at the school.

Conclusions

It is concluded that there is poor mental health at this residential school for the blind, as in most of the traits tested the mean percentile is far over on the unhealthy side of the scale, thus confirming the hypothesis. The researchers might pose two alternative causal possibilities. The personality difficulties might be due to the sensory loss itself, or to some peculiarity of the particular school. It should be noted in this connection that, in previous research (**4**), scores still decisively unhealthy but more closely approaching the norm were found in other residential schools, while scores clearly on the healthy side of the scale were found for the blind attending public schools. Thus some weight might be lent to the second alternative.

References

1. Adorno, T. W., Frenkel-Brunswik, Else, Levinson, D. J., & Sanford, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. Bernreuter, R. G. *Manual of personality inventory*. Stanford: Stanford Univer. Press, 1935.
3. Cutsforth, T. D. *The blind in school and society*. New York: D. Appleton, 1933.
4. Greenberg, H. *Some effects of segregated education on various aspects of the personality of those members of disadvantaged groups experiencing this form of education*. Unpublished doctoral dissertation.
5. Komisar, D., & MacDonnell, M. Gains in IQ for students attending a school for the blind, *Except. Child.*, 1955, **21**, 127–129.

Received March 28, 1957.

A PROCEDURE FOR MANAGING TEACHING-LEARNING SITUATIONS IN EDUCATIONAL PSYCHOLOGY

RAY H. SIMPSON

University of Illinois

Can the teacher of educational psychology seriously attempt to practice what is preached? Can the course in educational psychology, or any psychology course for that matter, give daily evidence to students that knowing psychological theory and principles really makes a desirable difference in the way teaching-learning situations are handled? In his own classes dare the teacher of psychology try to use what he knows in guiding the work of his own students? Most teachers of psychology have probably given some attention to questions such as these. The folowing description represents one attempt developed over a long period of years to answer these questions in the affirmative. It is not given as a model but rather as an approach which may be suggestive to others who are also concerned about possible ways of translating psychological theories and principles into practice in realistic teaching-learning situations.

Course Objectives

The general purpose of the course in educational psychology, as seen by the writer, is to prepare each student to do effective and satisfying professional work through the identification and application of psychological principles. Learning the attitudes and skills that will be needed for professional growth while teaching is emphasized; memorizing printed materials for test purposes is discouraged except as this can be shown to contribute to professional growth. Since the students in the current educational psychology course expect to teach, the attempt will be made to: (*a*) focus attention on the improvement of teaching and learning, (*b*) provide for cooperative consideration by the teacher and the learners in the class of the various processes involved in improving the teaching-learning situation, and (*c*) encourage the understanding and frequent use of varied methods of attacking professional problems in teaching-learning situations.

Course Activities

The writer's purposes in the educational psychology course are revealed in the outline of the major course activities, which are discussed in this paper: (*a*) identifying some key problems in teaching-learning situations, (*b*) selecting problems for initial study, (*c*) working toward problem solution, (*d*) keeping a student learning record, (*e*) self-evaluation, (*f*) finding and using resources, and (*g*) developing effective interpersonal relations.

With the guidance of the instructor, each student considers a great number of teaching-learning problems or situations. For example, consideration is given to such questions as these: On what basis should we group in this class? Should the total class group be used for all purposes? What use, if any, should be made of small groups? For what purposes might small groups be formed? Should out-of-class study groups be encouraged? What criteria should be used in marking in order to promote desired learning? These and similar questions relating to situations both in and out of the class vary with respect to implied educational objectives and methods. Some goals and activities are those of individuals, some of small groups, and some are those common to the whole class. In this class much attention is given to helping students develop sound goals and purposes which were not initially recognized.

Identifying Problems

In order to provide a sound basis for later individual study, each learner, during the first few days of the course, is encouraged to identify and list a number of professional problems that seem to be of concern to him. Those who have had teaching experience are encouraged to draw on this experience. Students without teaching experience may object that they do not know any professional problems. Actually, most students will have had a minimum of 13 years in teaching-learning situations. Through instructor guidance and some supervised observations of elementary and secondary school classrooms, the writer has found that undergraduate students without teaching experience show considerable ability in identifying teaching problems. However, direct experiences are not relied on exclusively.

Students are encouraged to use many types of resources in

widening their perception of the important issues now facing educators and of the problems they are likely to meet as teachers. Recommended resources for problem identification include: direct observation of classrooms taught by other teachers, educational psychology texts, other printed materials—particularly in educational and psychological magazines—other students and their difficulties, and other teachers and their problems. The following are problems illustrative of those listed by students:

How can I more adequately adapt my instruction to the individual differences in classes in ninth-grade general science? (Origin: own experience as a student in general science class which was not challenging to me.)

What techniques can a young and inexperienced administrator use to encourage older teachers to modify their practices? (Origin: talking with a recent college graduate who was somewhat discouraged.)

What are some of the guiding principles and procedures that can be used in working with slow learners in the classroom? (Origin: table of contents of an educational psychology text.)

What is good discipline and how can I promote it in my classes? (Origin: Experience as a student in classes. Also difficulties of a teacher in classes observed while I was a college sophomore.)

After the learners have compiled their lists, they are encouraged to get together in small groups to select the problems they consider most important. Sometimes this is done during class time if it is difficult for groups to meet outside of class. The goals here are to help each learner expand the scope of his perceived problems, to get the reaction of others, and to appreciate the kinds of problems which others see as important. It is believed that activities will help the students in their approach to the professional problems they will later encounter. Conventionally, of course, this is done, if it is done at all, almost entirely by the teacher.

Selecting Problems for Initial Study

After each student has identified a number of professional problems which he might consider studying (usually between 15 and 35), all are encouraged to help set up criteria which will aid in deciding which most deserve attention and study. The instructor usually suggests one or two questionable criteria and one or two better ones in order to get learners started, e.g.:

Questionable: What is easiest to teach or learn without regard to importance? What problem was studied first last year?

Better: What are the present and prospective needs of individual learners in the class? (It is hoped that the students will eventually encourage their pupils to consider similar criteria.) What kinds of needed psychological training or practice is each learner not likely to get in other phases of his educational preparation?

After hearing an example or two of each type, each student tries to list both desirable and questionable or overemphasized criteria. He is encouraged to give his own reactions, not to try to anticipate the instructor's attitudes. Subsequently, the ideas are pooled, and a list is put on the blackboard which involves the combined judgments of the class members. Students are encouraged to consider the significance of particular criteria in deciding what problems it is best to study.[1] The criteria which are judged to be the most valid are then used in deciding which problems will be studied intensively. The process just described helps each student develop basic purposes for his study and tends to de-emphasize excessive emphasis on working merely for a mark.

After the students have idendified a number of problems on the basis of the criteria previously listed, each, with instructor help, is encouraged to select two, three, or four problems and break each of them down into subproblems. An illustration of a problem with its subproblems that one student developed initially is the following:

How can I meet with a greater degree of adequacy the problem of individual differences in my classes in ninth-grade general science?

1. What are my specific purposes in teaching the course?
2. What are probable learner goals in taking the course and what are possible ways of determining these?
3. What plans should I consider making before the classes meet?
4. What are ways of determining a learner's science needs?
5. What are possible organizational setups, logical and/or psychological, which might be used in developing the course?
6. What are psychological strengths and weaknesses or limitations of various class procedures which other science teachers have proposed?
7. What procedures tend to develop critical thinking and the application of the scientific method to science problems in individuals of differing ability?
8. What are possible ways of evaluating course work and helping learners evaluate their work?

---

[1] For additional development of issues related to problem identification see: Ray H. Simpson, *Improving Teaching-Learning Processes*. New York: Longmans, Green, 1953, pp. 41–96.

9. What are ways of making group work more effective in promoting learning?
10. What kinds of resources will we be likely to need and how should these be obtained?

After the student has broken down two, three, or four problems into specific questions or subproblems—each problem with its subproblems constituting a possible working outline—the teacher and other learners in the group offer suggestions for improving them. From this group of two, three, or four tentative outlines, the student is in a position to decide which problem with its subproblems it would probably be best for him to study. After his decision has been made, the student may add, delete, restate, or reorder the subproblems. This refined statement of a major problem and its subproblems constitutes his basic organization or psychological frame of reference for early study. Several problems may be attacked by a student during the course of a semester.

Working Toward Problem Solution

After the learner has organized his subproblems into a tentative outline, he is in a position to start looking for solutions and to react critically to ideas which are gleaned from various resources. It is emphasized by the instructor that for any one educational problem there is no single "best answer" since the teacher and the learners differ significantly in their needs or abilities in different situations. For example, different approaches to discipline are usually called for in a rural neighborhood and in an urban slum district.

An attempt also is made to help the learner see that there are usualy arguments for and against each proposed solution, and he should become familiar with these arguments if he is to discuss the issues involved intelligently. Important in the understanding of the arguments is a consideration of the psychological principles or theories upon which each is based. A study of these principles and theories helps the learner build up a theoretical framework upon which he may make more valid judgments in other realistic problem situations.

Keeping a Student Learning Record

In the class in which a specific attempt is made to encourage the development of more effective learning processes, it is im-

portant that a rather complete record be kept by the student of his learning activities and ideas. Such a comprehensive record is useful from several standpoints. First, it enables the student to clarify his plans and ideas and the information acquired by writing them down in his own words. Second, it gives the teacher and other students an opportunity to examine and offer suggestions about the progress of his work in the course. Third, it provides an objective basis upon which, with the guidance of the teacher, he can evaluate the work that he has done to date and make plans for future activities.

The students are usually encouraged to keep a record of the following sort. Each should have an orientation section which includes a list of the problems which he has identified as important from his standpoint. This section should also include a breakdown of several problems which he has considered for initial study. In the orientation section of the learner's record there is also a list of the comments, criticisms, or suggestions which have been made by the instructor or by other students. The orientation section is usually followed by an outline of the main problem which he is attacking together with the subproblems under it.

Each learner keeps a resource sheet on which he lists, usually by number, all resources which he has found useful in connection with the course. After completing the resource sheet, the learner prepares a number of pages at the top of each of which he writes down one subproblem. As he finds possible solutions of his subproblems, he lists these on the appropriate pages. The origin of each possible solution is indicated by a number which refers to the location of the item in the bibliography or on the resource sheet.

Under each subproblem, after each group of ideas from a specific source, the learner writes his own critical comments and reactions to the ideas he has obtained from various sources. The primary purpose of encouraging learners to put down their own ideas is to develop critical thinking by each student. The various parts in the record, which are labeled "own reactions," include: one's own professional situation, relating theory to practice, relating practical suggestions to one or more theoretical considerations, pointing out weaknesses of proposals that have been given, suggesting types of situations under which the proposals might

work, and types of situations under which the proposals probably would not work. An attempt is made to urge students to go beyond such superficial comments as, "That looks good," or "I like that," or "it doesn't seem to me that there's anything in that," to expressions of reasons and a critical analysis of *why* the source or student takes a particular point of view.

Periodically, the learner's record is examined by the instructor and by other students in the class. The paramount purposes of such examinations on the part of the instructor are to get a general understanding of what each individual learner is doing, and to give comments and suggestions which will enable him to improve his work as he plans additional activities in the course. The peer evaluation provides helpful suggestions both to those examining a student record and to the student whose record is examined. The written evaluations and self-evaluations of records which students give are sometimes examined by the instructor. Through reciprocal student evaluations of records, the learner gets guided practice in offering help to other students.

Periodically, learners are encouraged to arrive at tentative conclusions with respect to particular subproblems that have been rather thoroughly investigated.

Self-Evaluation

Self-evaluation is conceived as making judgments concerning present and past activities and achievements *so that better plans may be made for future learning.* Thus conceived, self-evaluation is encouraged from the beginning of the course when the learner starts to identify problems of significance to him. Evaluation is a continuous process throughout the course, and at frequent intervals the student is encouraged to look over what he has done and, in the light of such an analysis, to plan future activities more effectively. At times the instructor and other students give guidance to the individual learner. One type of help in self-evaluation is to aid learners in setting up criteria which may be used in evaluating the individual record.

Below is a list of criteria developed by a subgroup in one class. The questions constituting the criteria can serve as one basis for self, peer, and instructor appraisal of student learning as shown in the record.

1. Are comments analytical, and do they show critical thought?
2. Have sufficient theoretical ideas from psychological resources been gathered?
3. Has a good balance of theoretical and applied ideas been maintained?
4. Have theoretical ideas been related to practice?
5. Have the theories underlying suggested practices been sufficiently examined?
6. Are there types of resources which should have been used more? If so, what are they?
7. Is the setup of the record flexible?
8. What do you think should have been done differently?
9. What types of things do you think you should stress more in future study? In what ways are these likely to improve learning activities?
10. Does it appear that your study, as you have developed it, is likely to improve your professional behavior?

In the course, several assumptions about evaluation are made, such as the following:

The primary purpose of evaluation should be to help the learner see where and how he can become more efficient in learning.

Learner goals and purposes in evaluation should be examined frequently and appropriate modifications should be made. Mental health is one of the paramount goals of self-evaluation. In the broad sense, no individual or group is likely long to remain mentally healthy unless efficiency in self-evaluation is developed. The teacher's own self-evaluation can constitute an important basis for teacher and learner improvement. Evaluation of learning activities *in process* is frequently more helpful in improving learning than evaluation of the *end* products of learning, since the former is the real basis for the latter and represents the place where improvement can best take place. Education of the types carried on in the course should not stop when a final grade has been obtained.

Each teacher and each learner should attempt, with appropriate help, to set up his own principles for evaluation, since it usually is, or at least should be, his own responsibility. All concerned should have a share in deciding the purposes and procedures of the evaluation. The student's share should increase as he gains additional facility. Autocratic evaluation basically presents an impossible job for the teacher, because of the number of students. Also, when attempted, it is likely to be uneven in quality, not only because of the magnitude of the job, but also because the learner himself *with sufficient guidance* is in the best position to evaluate his own

work. Individualization in education demands democratic evaluation.

Evaluation should not be thought of simply as the judgment of an end product, but rather as a continuing process which always goes on concomitantly with other phases of learning. It is a continuous appraisal and reappraisal for continuous reorientation of goals and activities. Nor should self-evaluation for learning purposes be one which allows students without guidance to do whatever they feel like doing. An important part of self-evaluation is helping the learner build up appropriate criteria to guide him in making various types of decisions. Evaluation should be made up of frequent analyses of past work, and should be oriented toward systematic plans for changes in future work.[2]

Finding and Using Resources.

Students are encouraged to examine different texts, source books, and other resources during the course. One principle to be kept in mind in getting and using resources is that those resources which are appropriate for one learner in a class may be most inappropriate for another learner in the same class. Some factors contributing to this principle are differences in needs, attitudes, skills, and basic abilities, and in what is considered usable material. The instructor has found that many students need much help in finding and using resources. Many students have been trained to depend almost exclusively on instructors and hence must be weaned gradually.

Purposes in using texts differ somewhat from those in conventional courses. Several texts are used rather than one. Each student uses different texts on a reciprocal basis with other students. One purpose is to help them develop the habit of referring to appropriate sources in problem situations. Informal studies of what teachers typically do in a professional problem situation indicate that such a habit is not currently common. A second purpose is to aid learners in practicing the skills necessary in getting help from educational psychology texts. Extensive use is made of tables of contents, introductory paragraphs in chapters, chapter summaries, book indexes, and other key parts of books. A third purpose of use of texts is to give practice in *critically* examining the theoretical

[2] For additional discussion and illustrations relating to self-evaluation see: Simpson, op. cit., pp. 173–229.

statements and practical suggestions in the texts. Initially, students may shy away from this, being loath to be critical of textbooks which they have come to regard almost as sacred cows. An attempt is made to help the learner see that through a critical analysis of the thinking of others his own thinking is sharpened.

Different texts are exchanged from time to time by individual learners and they attempt through use of the tables of contents, chapter introductions, chapter summaries, and indexes to use those parts of each book which seem to give the most theoretical or applied help on their particular problems.

The resources used by a student should depend upon the problems for study which he has selected, with the help of the teacher, and also should result partially from guidance given by the teacher during the course in the acquisition, selection, and use of resources. The instructor should have available a variety of bibliographies which are frequently used as "starters." Much utilized are *Psychological Abstracts* and *The Education Index.*

Students make considerable use of human resources, such as other students, teachers on the job, other instructors, and, of course, the instructor himself. They are also encouraged to use many other resources such as observations of classrooms in the public schools, and visits to youth groups. An annotated list of readily available films is given to each student and groups of learners are encouraged to view some films in the visual aids laboratory. Some of these films that small groups of students recommend highly are shown to the whole class at a later date.

Furthermore, they are encouraged to get practice in evaluating the ideas obtained from various resources and always to practice giving critical reactions to these ideas. Some instructors contend that, while the objective of such work is desirable, students are not capable of doing it. It is here maintained that competence is developed through guided practice in these activities.

Developing Effective Interpersonal Relations

One of the critical points in our culture seems to relate directly to interpersonal relations weaknesses. We believe in democracy and in some respects we are democratic. However, many school situations are characterized by essentially autocratic practices. A functional educational psychology must face the problems tied up with the question: Do educational and social psychological re-

searches indicate the need for a study and use of democratic interpersonal relations? Most of the evidence seems to suggest an affirmative answer. Therefore, since the course is considered a teaching-learning laboratory, attention is given to ways of studying and using democratic interpersonal relations. In most cases those who are affected by a particular decision should have an increasingly large share in making the judgments which lead to that decision.

Beginning with the first class meeting, an attempt is made to get participation by all students in planning and managing the work of the class. As the course progresses, the instructor attempts gradually to shift more and more social control and responsibility for learning activities from himself to the students. While early assignments are made largely by the instructor, these are of such a nature that the individual learner can adapt them to his needs and abilities. For example, an early assignment, in brief, might be:

> List specific professional problems related to teaching-learning situations which you have encountered as a student or teacher. Supplement this list by studying professional periodicals which take up problems in your field (for example, English teachers or prospective English teachers might use *The English Journal*). After each problem you list, indicate the situation or specific resource which suggested the problem to you.

In one sense all students have the same assignment; however, it is set up in such a way that each learner can adapt it to his background, interests, and needs. If students had had a more desirable type of training in individual responsibility for learning activities, it is quite possible that a more rapid start could be made in giving them primary responsibility for setting up their own learning activities. However, in view of the fact that learners most often come to the class expecting to be told what to do, when to do it, and how to do it, it is necessary, in line with principles of readiness, to work with the learner where he is and get him started off in assuming responsibility gradually.

The following are some of the approaches which have been found useful in establishing effective communication between the students and the instructor and among the different class members: An advisory committee from the class may help the instructor plan the framework for in-class and out-of-class activities. This group may be selected in a variety of ways, any one of which may be the best in a particular situation. Sometimes the

advisory group is made up of volunteers. In other classes it has seemed advisable for the instructor to select representatives from various subject-matter or teaching-level groups if the class believes that would be most desirable. In some classes, where small groups within the class are formed early and students have become acquainted, it is possible to have each group select a representative who serves on the advisory committee for a definite length of time. In other classes, where students have known each other, or each student has known several others, it is sometimes desirable to elect representatives. When nominations and elections are used, such questions as these are sometimes raised: What kinds of students should we try to elect? For example, should only the best students be elected or should we elect some who are weak (as determined by self-evaluation, teacher evaluation, or other-person evaluation) in this type of activity so that they can get practice? Students usually decide it is wise to have some of both types. Typically, it is probably desirable to discuss the method of choosing the advisory group so that the students may get some experience in this kind of learning activity.

In large classes, frequent conferences between the instructor and individual students are not possible. As a substitute, a written communication system can be introduced. On a cover sheet of the learner's record, he can note down his problems and difficulties. The teacher can collect a certain number of learner records with attached comment sheets each day or each week. Between class periods the instructor examines them and returns the records with his comments on the cover sheet at the next class period. These may note learning strengths shown by the record, weaknesses in organization, or other criticisms, and may suggest additional approaches or resources. This arrangement permits the teacher to look over the learning record of each student, see what he is doing, and offer constructive suggestions to the learner. It also permits the student to indicate to the instructor what he believes to be his primary difficulties. In addition to regularly scheduled inspections, it is debatable whether it may sometimes be desirable to have unscheduled inspections by the teacher. They permit the teacher to see work in process rather than a record "fancied up" for the purpose, and so encourage the student to get help on a voluntary basis.

Time is occasionally given in helping learners develop in the

various roles that a mature group member uses at various times in classes and small groups. Roles include such ones as are involved in giving leadership to an unstructured group. In some classes a bibliography which deals with individual roles in a group and on problems of "growing a group" is given to students. Individual members study various items of the bibliography and also look up other references of a similar sort and then report back either to the class or to small groups in which they may be working.

It must be recognized that the teacher is in a key position as major class leader for guiding group activities toward democratic values and sound procedures. The democratic teacher is able to take upon himself a great variety of roles in the group. In any particular situation, he attempts to assume the role which promises to be of most value to learners. This will sometimes be a good listening role. To promote a situation in which learners can practice different healthy interpersonal roles, it is necessary for the teacher to depart from some roles traditionally played and assume many different roles at different times. For example, the teacher probably should not be thought of as knowing all the answers to the realistic problems that can come up in the course of discussion. The teacher's job, rather, is one of knowing a lot of answers and also of knowing how to help learners develop ways of assuming the responsibility for getting and analyzing possible answers themselves.

Typically, the learner's concept of his responsibility for class operation needs to be changed. The idea is current that the teacher is almost solely responsible for the whole class and for individual and small group success. Added learner responsibility would seem to be essential for significant improvement in large and small group activities as well as in individual learning.

One aspect of interstudent communication is the setting up of groups. These groups may be set up for different purposes. Some of the purposes which have been found significant are:

To answer each other's questions with respect to the operation of the course and the development of individual learning activities.

To offer critical and constructive suggestions on the learning record of other members of the group.

To develop, as a group, various aspects of a particular problem. For example, students in one group might be working on the question: What are some of the ways of diagnosing disciplinary problems and improving

them? One student might work on the matter of improving assignment making, another on the ways of getting and using appropriate resources for class activities, and another on the issue of getting democratic learner participation in class activities.

Some consideration has already been given to evaluation. It is important in this connection that both the teacher and the learner give consideration to ways of evaluating their individual behaviors in interpersonal situations and in planning improvement in such behaviors.

Class cooperation is sought not only in developing plans for in-class and out-of-class activities, but also in setting up evaluative criteria which will be used in promoting improvement in the class, and as a basis for marking where marking is either required by the system in which the class is operating or is felt desirable by both the teacher and students. When marking becomes a dual responsibility of the teacher and learners, it is more likely that an attempt will be made to align the basis of marking with the types of activities which will promote profitable learning. This means that the teacher and students must together examine and re-examine the primary purposes of the course and the degree to which these purposes are being achieved.

In the light of such purposes, the teacher and the students can cooperatively set up criteria on which the marks will be based. Such criteria may differ quite significantly from one course to another, depending on its purposes. For example, if it were a course in appreciation of literature, the primary criterion on which marking would be based might be: Has there been improvement in the appreciation of literature defined in behavioral terms during the semester or year and, if so, to what degree? On the other hand, if it is a professional course such as educational psychology or the psychology of learning for teachers, a major criterion might be: To what level of professional development has this learner grown by the end of the period of study—if the mark is designed to indicate level of professional competency.

Summary

The following points represent what the writer believes to be crucial aspects of the course which has just been described.

Continually, attempts are made by the instructor to have each member of the class voluntarily bring his intelligence to bear upon

the total problem of utilizing psychological resources in the identification and the solution of significant professional problems—significant particularly to the individual learner. Generally this process involves setting up goals and purposes, deciding on courses of actions, and attempting a continuous evaluation of the work as it proceeds.

The course for an individual learner is not rigidly set up in advance. In effect, the learner produces a custom-tailored set of learning experiences with the guidance of the teacher.

Goals, study procedures, class procedures, and other factors in the management of teaching-learning situations are kept flexible enough to give considerable attention to the development of learner judgment.

Learning for the student consists of a combination of individual, small-group, and whole-class activities which are planned by those most concerned.

The educational psychology course is, itself, used as a teaching-learning laboratory where informal tryouts of some proposed solutions to educational problems can be made.

Received April 23, 1957.

SEX DIFFERENCES AMONG TEACHERS IN THE USE OF THE CRITERIA OF ADJUSTMENT[1]

HARRY BEILIN AND EMMY WERNER

University of Minnesota

The criteria employed by teachers in evaluating behavior problems of children has been a matter of much concern since the classic study by Wickman reported in 1928 (**25**). In that study, the attitudes of teachers and "mental hygienists" were contrasted in regard to a list of children's undesirable traits and behavior. The results indicated greater teacher emphasis upon classroom management, learning, and authority problems; clinicians placing greater emphasis upon withdrawing and other "unsocial" personality characteristics.

The report of these findings is regarded as having had a profound effect upon teacher education programs (**6**) with directly measurable changes in teachers' attitudes evident as early as five years after publication of the study (**1**). To say the Wickman study itself was responsible for the reversal of emphasis in teacher education would not do justice to the tenor of the times which the spirit of the report reflected: namely, the growing awareness among many groups of the psychological approach to the understanding of human behavior. This may account for the astonishment of some observers (e.g., **6**) that the Wickman study should have been so widely quoted and uncritically accepted when it embodied some basic flaws.

The limitations of the investigation have been noted as involving three general problems: methodology, the adequacy of the criterion, and the consistency of behavior related to adjustment. The problem of methodology is not independent of the others. The criticisms are in essence these:

1. *Methodology.* The directions given teachers and clinicians

[1] This study was supported in part by a research grant (M-690) from the National Institute of Mental Health, and in part by funds from the Institute of Child Welfare, University of Minnesota. This is one of a series of papers of the Nobles County Project, J. E. Anderson and D. B. Harris, Principal Investigators.

were different in the time to respond and the emphasis upon present seriousness of behavior as against seriousness for future adjustment. The differences in directions should affect the comparability of results for both groups.

Replications of the study and related investigations have employed the same directions for teacher and clinician (or other) groups, using the original directions given clinicians (**7, 12, 20**) or those to both groups (**4, 18, 20**). There have been replications (e.g., 25 years later) retaining the original design (**20**) as well as introducing a variety of modifications (**13**) including changes in the time-seriousness orientation (**7, 9**). The results, in the main, indicate that the directions given *S*'s do account to some extent for the differences obtained between groups, but disparities still emerge. With the passage of time, however, there has been a tendency for teachers' criteria to more closely approximate those of clinicians.

2. *The criterion problem.* The second issue concerns the choice of clinicians as a group against which to refer teachers' attitudes toward children. Watson (**23**) has pointed out that there is no reason to suppose clinicians to be "correct" and teachers not, rather than vice versa. Wickman is criticized for not even considering the possibility. Other investigators (**6, 10, 19**) have dealt with this issue although not with the attention it deserves. It has been our thesis (**2**) that the criteria of adjustment are determined, at least in part, by the status and role of the person who evaluates the adjustment of others, and holding to a set of criteria is not inherently more correct for one group than another. The adequacy of criteria (of adjustment) will depend upon some independent standards related to the values and goals of those making use of the criteria. Since teachers may have very different goals from clinicians regarding the behavior of children, it may not be appropriate to impose the clinicians' goals upon the teacher.

3. *Continuity of behavior related to adjustment.* The last major limitation of the original study was its too ready acceptance of a direct relationship of withdrawing behavior in childhood to maladjustment in adulthood (**14, 23, 25**). No adequate data were available to support this assumption in 1928, and one may venture to say no definitive data are available now to substantiate it. As we have observed elsewhere (**3**), a too ready acceptance of this

assumption has tended to hamper success in studies of the prediction of adjustment from childhood to adulthood.

In the present study, we are concerned with the problem of the criterion. In exploring further the relation of role and status to the criteria of adjustment, we are focusing on the effect of sex roles. It has been suggested by Davis and McGinnis (**6**) that the differences between teachers and clinicians may have been due to sex differences (since in 1928 teachers were more likely to have been women; clinicians more likely to have been men[2]).

Procedure

In the present investigation, verbal descriptions from men and women teachers of the behavior of adolescents judged by them to be well or poorly adjusted were obtained. This procedure differs from Wickman's in that teachers are not provided with a set of behavior descriptions but offer their own. The present approach or modifications of it have been employed by others in investigating teacher attitudes (**5, 13, 16, 17, 26**). The present study differs from most all those cited in that the same amount of emphasis is placed upon the criteria of adjustment as on maladjustment.

Thirty-nine teachers (17 women, 22 men) of the high school of a Minnesota community of 8,000 were interviewed as part of a series of studies in the prediction of adjustment. For the purpose of the present investigation, each teacher was asked to nominate the three best and the three poorest adjusted students (regardless of sex) in that school. The selection of high school teachers as evaluators purposefully limited the age range of the persons they evaluated. The teachers were then specifically asked: "Considering the three people whom you have nominated as best adjusted, what is there about them as a group or individually that makes them stand out as best adjusted?" (The same was then asked with regard to the persons nominated as poorest adjusted.)

The responses of the teacher-evaluators were recorded, content analyzed, and the frequencies of responses for each category tabulated according to the sex of the evaluator. Percentages were computed and are presented in Tables 1 and 2. A critical ratio formula

---

[2] Thompson (**22**) reports that women teachers' rankings agree somewhat better with those of child psychologists than do the rankings by men teachers (data however are not presented). Epstein's study (**8**) is also concerned with sex differences in ranking behavior problems of children.

TABLE 1

DIFFERENCES BETWEEN MEN AND WOMEN HIGH SCHOOL TEACHERS IN MENTION OF CRITERIA OF GOOD ADJUSTMENT

Criteria	Female (N = 17) %	Male (N = 22) %
1.0 *Emotional-personal* (total)	82.3	100.0
1.1 mature, good judgment, sensible, sound	11.7	36.4*
1.2 pleasant, likeable, sweet sunny disposition, good personality, personality	23.5	36.4
1.3 adaptable	23.5	9.1
1.4 general attitude is good, healthy attitude toward marriage, sex, surroundings, other persons	0.0	18.2
1.5 are secure, don't show off, don't complain, not self-conscious	5.8	27.3*
1.6 well-integrated, emotionally stable, well-balanced, calm, not tense, poised	29.4	40.9
1.7 accepts criticism, has insight into self	23.5	9.1
1.8 vivacious, lively, enthusiastic, energetic, glowing	17.6	4.5
1.9 happy	11.7	0.0
2.0 *Character-control* (total)	94.1	72.7*
2.1 good manners, show respect for authority, attentive, courteous, polite, no discipline problems, well-behaved, accept mores of society	29.4	45.5
2.2 takes responsibility, accepts responsibility, assumes responsibility beyond call of duty, trustworthy, dependable, reliable	58.8	45.4
2.3 honest, sincere, straightforward, genuine	5.8	0.0
2.4 are not led, have principles, mind of their own, independent, self-sufficient, real strong individualist	29.4	9.1
2.5 keep themselves decent, have not wandered, good character, don't get into difficulties, moral conduct o.k., tries hard to follow straight and narrow path	0.0	9.1
2.6 humble, modest	23.5	4.5*
2.7 unselfish, altruistic, wants to serve fellow men, do good for others	5.8	0.0
2.8 good at getting things done under trying circumstances, handles difficulties well	5.8	18.2
3.0 *Social-interpersonal* (total)	88.2	77.3
3.1 charming, meets people well, likes other people to be comfortable, ability to carry on conversation, good mixer, friendly, extrovert, fits in well, gets along well, hosts of friends, well liked, popular	70.6	68.2

TABLE 1—*Continued*

Criteria	Female ($N = 17$) %	Male ($N = 22$) %
3.2 leader	35.3	27.3
3.3 cooperative, willing	11.7	13.6
3.4 aggressive, lots of guts with people	0.0	9.1
3.5 participation in extracurricular activities, community activities	23.5	9.1
4.0 *Achievement in general and specific cultural roles* (total)	88.2	72.7
4.1 ambition to make good, wants to get ahead, doesn't let opportunities go by, are goal-oriented, know exactly what they want, planful, lots of determination, drive, initiative, sticks to things and accomplishes them	58.8	36.4
4.2 good students, use their ability, like school, high scholastic standing, get good marks	41.2	18.2
4.3 not necessarily the best or good students but they try	11.7	9.1
4.4 does good job, carries on job very successfully, terrific worker, tends to business, doesn't loaf, works well independently, industrious, efficient at job	17.6	22.7
5.0 *Abilities and interests* (total)	52.9	45.4
5.1 all round person, variety of interests	23.5	18.2
5.2 have ideas, think for themselves, logical thinkers, inventive, creative	11.7	13.6
5.3 bright, alert, capable, superior, brilliant	29.4	13.6
6.0 *Appearance* (total) attractive, very gracious, very fine dresser, carries herself well, neat appearance	5.8	4.5

* Significant at the 10% level of significance.

for independent proportions (**11,** p. 60) was used to determine the significance of the differences in response by sex to each category. The tables, then, indicate the proportion of each teacher group (male and female) that employed each criterion and where significant differences appear. The specific criteria are grouped into general areas: *social-interpersonal,* statements reflecting the relationship between the *S* and others; *emotional-personality,* statements which indicate adequacies or inadequacies in adjustment in which the social reference is lacking or minimal; *character-control,* statements which indicate the extent or adequacy with

TABLE 2

DIFFERENCES BETWEEN MEN AND WOMEN HIGH SCHOOL TEACHERS IN MENTION OF CRITERIA OF POOR ADJUSTMENT

Criteria	Female (N = 17) %	Male (N = 22) %
1.0 *Emotional-personal* (total)	82.4	81.8
1.1 seeks attention, shows off, wants to be something he isn't, bluff exterior only	23.5	27.3
1.2 persecution complex, suspicious of adults, feels people are picking on him	17.7	27.3
1.3 dissatisfied with things as they are, looking for something they haven't got	5.8	4.5
1.4 emotionally unstable	41.2	18.2
1.4.1 unstable, inconsistent in moods	11.7	0.0
1.4.2 violent temper, bad temper, uncontrolled, purely emotional, hot-headed, high-strung, quick-tempered	5.8	9.1
1.4.3 nervous, can't settle down, restless, moves around, tense, lives in constant fear	23.5	9.1
1.5 resents suggestions from others, reacts poorly to criticism	0.0	18.2
1.6 stands out as different from others, a queer duck a little bit odd, strange	11.7	9.1
1.7 insecure, lacks confidence in self, afraid to say wrong thing, seeks praise, ill at ease, always full of alibi	11.7	22.7
1.8 stutters when excited, one complaint after another	5.9	9.1
1.9 immature, acts like baby, kind of young	11.7	4.5
1.10 unhappy, gets no enjoyment from things average people do	11.7	4.5
1.11 smiles inappropriately, silly sneer on face	5.9	4.5
1.12 makes fool of self, others feel sorry for him	0.0	4.5
1.13 daydreams, reads escapist books	0.0	4.5
1.14 sissy, feminine characteristics, possible sex problems	0.0	4.5
1.15 no spark	5.9	0.0
1.16 sadistic, hurts animals	5.9	0.0
2.0 *Character-control* (total)	82.4	63.6
2.1 unreliable, not dependable, lacks sense of responsibility, untrustworthy, tatler, lies	0.0	27.3*
2.2 hard kid to handle, very trying, lots of disciplinary problems, negativistic, hostile to authority, defiant, cocky, smart alecky, belligerent attitude, in conflict with mores, rebels, poor attitude, chip on shoulder, wise guy, trouble maker, hostile, gets into trouble in school, mischief maker, poor conduct	58.8	31.8*

TABLE 2—*Continued*

Criteria	Female ($N = 17$) %	Male ($N = 22$) %
2.3 boisterous, foolishness and joking all the time, conduct themselves poorly in public, loud, noisy, discourteous, coarse no manners	35.3	22.7
2.4 very dependent, not self reliant, can't stand on own two feet, can't make decisions, mamma's boy, conformist, conforms precisely to accepted pattern	5.9	18.2
2.5 runs with wrong crowd, easily led by others in crime	0.0	4.5
2.6 too independent, goes own way, grows up as sees fit	11.7	0.0
3.0 *Social-interpersonal* (total)	70.6	63.6
3.1 lone wolves, withdrawn, unresponsive, don't communicate, shy, timid, very quiet, passive, aloof reserved, not comfortable with people, typical wallflower, mousy, standoffish	35.3	45.5
3.2 lack of adjustment to social groups, not liked, no friends, not accepted in group, not accepted by others, get along poorly, don't attract people, not friendly with people	41.2	35.3
3.3 are uncooperative with peers	0.0	9.1
3.4 self-centered, snobs, envious of others, critical of others, overbearing	17.1	4.5
3.5 overly aggressive, domineering, too bossy, bully	5.9	4.5
4.0 *Lack of achievement* (total)	47.1	63.6
4.1 no goals, no desire to get ahead, no purposefulness, doesn't know what she wants, just rides along, just rolls along, no get up and go, no ambition to get anything done, lazy, no drive for self-improvement, can't buckle down, lack effort, don't apply themselves, never stick to anything	23.5	45.5
4.2 failure at school, contributed little at school, repeating grade, no interest in school, poor student	23.5	27.7
4.3 can't read, can't spell, trouble in expressing himself	11.7	9.1
5.0 *Lack of ability and interest* (total)	41.2	18.2
5.1 lack of ability, low scholastic ability, not bright, slow to learn, doesn't seem to understand anything, not much on the ball	29.4	13.6
5.2 bored, uninterested, dull, inattentive	11.8	9.1
6.0 *Appearance* (total)	17.6	13.6
6.1 physical handicap, long-term illess, physically	11.7	9.1
6.2 not too attractive, slovenly appearance, dirty, personal habits poor, nothing glamorous, dresses inappropriately	5.8	9.1

* Significant of the 10% level of significance.

which the individual has incorporated social controls into his repertory and the traits developed in the process of doing so; *abilities-capacities,* statements referring to *S*'s abilities, capacities, and skills; *achievement,* statements which indicate adequacy or inadequacy with regard to achievement generally or to specific cultural roles, e.g., school, job; *appearance,* statements concerning personal appearance.

Results and Discussion

The reliability of the categorization process was determined for this series of studies employing two psychologically trained raters as classifiers. The percentage agreement for the classification of specific criteria was 85%. Agreement with regard to area classification was 91%.

For six categories (out of 83) there appear significant sex differences in frequency of response. Since the 10% level of significance has been taken as the cutoff point, it is apparent that the number of significantly differentiating items could have occurred by chance alone. The fact that the pattern among the differences is consistent with other results that have been obtained, however, suggests that these particular differences may be meaningful. The most significant observation to be made concerning the results, however, is that male and female teachers are to a high degree consistent in their employment of the criteria of adjustment.

The differences which do emerge suggest that male high school teachers are more inclined to value maturity, good judgment, dependability and trustworthiness. They tend, as well, to characterize the best adjusted as secure persons and lacking in self-consciousness.

Women high school teachers, on the other hand, place greater emphasis on those items classified together as "good character" items, particularly humility and modesty. In addition, and probably most characteristically, women teachers place more stress upon negativism, hostility to authority (the teacher), discipline problems, etc. as criteria of poor adjustment.

As already indicated, in all other respects the criteria which are employed in describing the least and the best adjusted high school students are essentially the same for both men and women teachers. In assessing the relative frequency of the criterion variables (by ranking) for each sex by *area,* some of the findings already

indicated receive further support. The most important area mentioned for both good and poor adjustment by women teachers are the "character-control" (essentially, classroom management) items. For the male teachers, however, the "emotional-personality" items receive prime mention. Achievement and ability variables are mentioned quite frequently (as much as 80%) by the high school teachers in characterizing those best or poorest adjusted.

These results tend to lend support to the hypothesis that there may be some sex differences in adjustment criteria and in the direction suggested by Davis and McGinnis, namely, greater emphasis by men on the emotional-personality criteria; greater emphasis by women on the character-classroom management criteria. Considering the possibility these results could have occured by chance alone, however, makes these findings only suggestive.

Summary and Conclusions

Considerable agreement was found between men and women teachers in the criteria used to characterize best and poorest adjusted high school youth.

The criteria which are used significantly more by men teachers are maturity, good judgment, dependability, trustworthiness, lack of self-consciousness, and being secure as a person.

The criteria employed more characteristically by women teachers are humility and modesty. They also place greater stress in regard to poor adjustment upon negativism, hostility to authority, discipline problems, and getting into trouble in school.

In general, the "character-control" items as a group are more often mentioned as criteria by women teachers, "emotional-personality" items as a group by the men teachers.

Intellectual and achievement variables play an important role in the assessments teachers make of the adjustment of young people.

These results are considered in light of a sex difference hypothesis concerning the differences between clinicians and teachers obtained in the Wickman study.

References

1. Bain, W. E. A study of the attitudes of teachers toward behavior problems. *Child Developm.* 1934, **5**, 19–35.

2. Beilin, H. Effect of social (occupational) role and age upon the criteria of mental health. *J. soc. Psychol,* in press.
3. Beilin, H. The prediction of adjustment over a four year interval. *J. clin. Psychol,* in press.
4. Bowles, H. H. A study of nurses attitudes toward behavior problems in children under hospital care. *Child Developm.* 1937, **8,** 282–288.
5. Boynton, P. L., & McGaw, B. H. The characteristics of problem children. *J. juv. Res.* 1934, **18,** 215–22.
6. Davis, E. A., & McGinnis, E. *Parent Education.* Minneapolis: Univer. Minnesota Press, 1939.
7. Ellis, D. B., & Miller, L. W. Teachers' attitudes and child behavior problems. *J. educ. Pschol.,* 1936, **27,** 501–511.
8. Epstein, L. J. An analysis of teachers' judgments of problem children. *J. genet. Psychol.,* 1941, **59,** 101–107.
9. Laycock, S. R. Teachers' reactions to maladjustments of school children. *Brit. J. educ. Psychol.,* 1934, **4,** 11–29.
10. McClethanon, R. H. Teachers and parents study childrens' behaviors. *J. educ. Sociol.,* 1934, **12,** 325–33.
11. McNemar, D. *Psychological Statistics.* New York: John Wiley, 1955.
12. Mitchell, J. C. A study of teachers' and mental hygienists' ratings of certain behavior problems of children. *J. educ. Res.,* 1942, **36,** 292–307.
13. Peck, L. Teacher reports on the problem of unadjusted school children. *J. educ. Psychol.,* 1935, **26,** 123–138.
14. Preston, G. H., & Shepler, W. McL. A study of the problems of normal children. *Amer. J. Orthopsychiat.,* 1931, **1,** 245–256.
15. Schrupp, M. H., & Gjerde, C. M. Teacher growth in attitudes toward behavior problems of children. *J. educ. Psychol.,* 1953, **44,** 203–214.
16. Seidman, J. M., & Knapp, L. B. Teacher likes and dislikes of student behavior and student perceptions of these behaviors. *J. educ. Res.,* 1953, **47,** 143–149.
17. Snyder, L. M. The problem child in Jersey City elementary schools. *J. educ. Sociol.,* 1934, **7,** 343–352.
18. Sparks, J. N. Teachers' attitudes toward the behavior problems of children. *J. educ. Psychol.,* 1952, **43,** 284–291.
19. Stewart, N. Teachers' concepts of "behavior problems" in *Growing points in educational research.* Wash. D. C. Amer. Educ. Res. Ass., 1949.
20. Stouffer, G. A. W., Jr. Behavior problems of children as viewed by teachers and mental hygienists. *Ment. Hyg.,* 1952, **36,** 271–285.
21. Stouffer, G. A. W., Jr. The attitude of secondary school teachers toward certain behavior problems of children. *Sch. Rev.,* 1956, **64,** 358–362.
22. Thompson, C. E. The attitudes of various groups toward problems of behavior. *J. abnorm. soc. Psychol.,* 1940, **35,** 120–125.
23. Watson, G. A critical note on two attitude scales. *Ment. Hyg.,* 1933, **17,** 59–64.
24. Wickman, E. K. *Children's behavior and teachers' attitudes.* New York: Commonwealth Fund, 1928.

25. Witmer, H. L. Childhood personality and parent-child relationships of dementia praecox and manic depressive patients. *Smith Coll. Stud. Soc. Wk.*, 1934, **4**, 287–378.
26. Yourman, J. Children identified by their teachers as problems. *J. educ. Sociol.*, 1932, **5**, 334–43.

Received April 24, 1957.

CHRONOLOGICAL DIVISIONS OF THE LIFE SPAN

HORACE B. ENGLISH

Ohio State University

Most developmental psychologists object to the notion of distinct "stages" of development. Yet it has been almost universally found convenient to describe individuals by such terms as child, adolescent, adult. To find, for the *Dictionary of Psychological and Psychoanalytic Terms,* how these should be defined, I sent a simple open-ended questionnaire to the fellows of the APA Division on Developmental Psychology, asking them to state the chronological limits applicable for these and similar terms. The freedom of expression encouraged did not make for simplicity of tabulation; but, as hoped, a fairly clearcut consensus as to usage was revealed. One hundred and sixteen questionnaires were sent out, 74 were returned.

Table 1 sets forth the limits which best meet the consensus for each named temporal division, and the percentage of replies favoring those limits. In every case there is a mode in the distribution, but in a few cases a secondary mode shows that a divergent principle of categorization is being employed by a minority. In most instances, however, it is the adjacent categories which are next most favored, suggesting that the minority would easily accept the modal usage.

The numerous comments are interesting. "Neonate" is taken as a complete synonym for "Newborn," and many favor its abandonment as pedantic and unnecessary. (Amen to that, and the same for a lot of similar Greek and Latin neologisms!) The "Gang Age" does not belong in the list; it refers to a "stage" in development, if it means anything, which many respondents doubt. There is no consensus as to when it occurs and little desire to pin it down. "Puberty" is also rejected from this list; it is a physiological condition, not an arbitrary division of the life span to be defined by usage or consensus. Unfortunately, I assumed that adulthood and maturity would be used synonymously and did not list the former term. A few respondents, however, seem to need a period of "early

TABLE 1
DIVISIONS OF THE LIFE SPAN

Infancy		Childhood		
Fr. Birth to 11.9 Mos. (82%) (46%)		1 Yr. to 11.9 Yrs. (33% (41%)		
Newborn (Neonate) Birth to 0.9 Mo. (100%) (42%)		Early Childhood (young child) 1 Yr. to 5.9 Yrs. (30%) (50%)	Mid-Childhood 6 to 9.9 Yrs. (22%) (25%)	Late Childhood (pre-adolescence) 10 to 11.9 Yrs. (35%) (50%)

Adolescence			Maturity	Old Age
12 Yrs. to 20.9 Yrs. (60%) (42%)			21 to 65 Yrs. (32%) (24%)	65 Up (47%)
Early Adolescence 12 to 13.9 Yrs. (50%) (44%)	Mid-Adolescence 14 to 15.9 Yrs. (50%) (45%)	Late Adolescence 16 to 20.9 Yrs. (35%) (45%)		

adulthood" between adolescence and maturity; they then define the latter as beginning at 30 or even 40. Most, however, are content to begin maturity with the twenties. Perhaps the gerontologists would be inclined to suggest a later age than 65 as the conventional beginning of old age.

The most significant comments, however, were those expressing reluctance to accept a chronological bounding of any of these periods. Quite rightly, they insist that no exact psychological meaning can be given to most of these terms. That, it seems to me, is precisely why a merely formal definition in chronological terms is best. If, for example, one defines adolescence as beginning with puberty, one seems to imply a specific pattern of psychological behaviors "characteristic" of the stage. Such a clearcut stage probably does not exist. Discussions based on such a usage mislead the layman and do not help the professional.

Let us frankly admit that all the distinctions are arbitrary. There is not one of the life periods which could not perfectly well

have its limits shifted forward or backward. Is a "newborn" an infant of less than 10 hours, of 10 days, of 14 days, or of what? It doesn't matter which we say except that it is better for us all to say the same thing. There should be strong reasons given if one elects to deviate from the consensus of expert usage here set forth.

Received May 9, 1957.

A COMPARISON OF READING COMPREHENSION SCORES OBTAINED BEFORE AND AFTER A TIME ANNOUNCEMENT

DESMOND L. COOK

Purdue University[1]

One of the most commonly used methods for measuring the rate of reading is to ask the reader to indicate, at a given time signal, the point he has reached in a given reading selection or passage. Rate scores obtained by this method are usually expressed in terms of the number of words read per unit of time, or some similar ratio.

A variation of the above technique is employed on the reading comprehension test given to all new students entering the University of Iowa. This test consists of a series of 10 reading passages, each followed by several choice type questions designed to measure various comprehension skills and abilities. The total number of items is 80, and the time limit for the test is 90 minutes. At the start of the test, examinees are informed that at the end of 45 minutes (or when one-half of the time has passed) they will be asked to indicate on the answer sheet the last item marked. The item indicated serves as a measure of reading rate and is so reported.

The validity of such a procedure may be open to question since such a score would not be a measure of pure reading rate. It is a measure of the time spent not only in reading the passage but in reading and answering the questions following each passage. Thus it is possible for two students to mark the same item but yet have actually completed different amounts of the test. For example, student A may have just completed marking item 42 at the time announcement, but student B has answered item 42 and gone on to read the next passage. Under the procedure of obtaining rate scores, both students would mark item 42 on the answer sheet. Nevertheless, it has been demonstrated to be an adequate measure of the rate at which students read (**1, 8**).

It seems logical to believe that those examinees who had been reading slowly prior to the time announcement would speed up

[1] The data for this study were obtained when the author was Associate Director, Examinations Service, State University of Iowa, Iowa City, Iowa.

in order to complete the test within the set time limit. Such a speed-up could result in lower comprehension scores for this group due to the fact that not as much time could be given to careful reading of the passages and answering of the questions. The faster readers meanwhile would be further along in the test and would be less concerned about speeding up to finish the test. They would have sufficient time to finish the test, and, consequently, the comprehension score would not be affected to the same degree as for the slow readers. It is even conceivable that the faster readers might improve their scores since more time could be devoted to the remaining items.

The relationship between the rate of reading and comprehension has been the subject of a rather large number of studies. Traxler (**10**), Traxler and Townsend (**11, 12**), Tinker (**9**), and Stroud (**6**) present rather comprehensive summaries evaluating the research on this problem. Stroud reports that the average correlation between speed and comprehension is around .40 when timed tests are used. As Traxler points out, this condition is to be expected if both measures are secured from the same timed test. Using untimed tests, Stroud and Henderson (**7**) found that for elementary school children there was virtually no correlation between rate and quality of reading. Using both timed and untimed tests with college students, Preston and Botel (**4**) report a correlation of .20 for untimed tests and .48 for timed tests, the latter correlation being significant. Usually these studies relating rate to comprehension have been concerned only with the relationship between these two variables and have not reported on the accuracy of comprehension involved.

Specific study of the accuracy of comprehension scores obtained before and after time announcement used to secure a rate score on a timed test appears not to have been investigated. Flanagan (**2**), Anderson and Tinker (**1**), Tinker (**8**), and Robinson and McCollum (**5**) have approached the problem by noting comprehension accuracy in set time limits, but they do not report comparative results. Flanagan studied the relationship between rate and comprehension using twelfth-grade students trained to read at varying rates. Comprehension was noted to decrease slightly as the students went from a slow to a medium rate and much more from the medium to a rapid rate. When the students were divided into three groups on the basis of the slow rate, each group showed

about the same decrease in comprehension scores from medium to rapid rates. The results were obtained, however, using scores for the full testing time, and no evidence was reported on relation of accuracy for items attempted. Anderson and Tinker report the results of a study concerned in part with the number of items attempted and scores on these items in a standard time (i.e., part time) for various parts of the Iowa Silent Reading Test. A correlation of .88 was found for the composite score with regard to scores in standard time with items attempted in the same time. The tests were administered individually to 110 college sophomores. No comparison of comprehension accuracy before and after the standard time announcement was reported. Using tests of varying reading difficulty, Tinker found that comprehension accuracy in a standard time varied from no difference to an accuracy of response of 60–70% for the most difficult test. In this study, the standard time limits varied from test to test and ran from 1¾ minutes to 36 minutes with students being allowed unlimited time to finish the tests. Again no results were presented regarding the comparative accuracy of response before and after the standard time announcement. A study of relationship between accuracy of response to number of items attempted on the Paragraph Comprehension part of the Iowa Silent Reading Test for 37 good and 33 poor readers was made by Robinson and McCollum. The groups were selected on the basis of their comprehension scores on the total test. It was found that the good readers were usually superior to the poor readers in both rate and accuracy. Essentially the same results were obtained when the groups were given the Van Wagenen "A" and "Alpha" tests which were untimed. The authors concluded that efficiency in rate of reading was the more influential factor in reading test superiority. As was the case in the other studies cited above, the comprehension accuracy results were reported for items attempted in either a standard or total time, or the accuracy of response on untimed tests, with no comparison of accuracy before and after a standard time announcement on a timed test.

Purpose

The purpose of this investigation is to compare comprehension scores obtained before and after a time announcement for groups of slow and fast readers. The specific hypotheses to be tested are

that the mean proportion of items correct before a time announcement is the same as the mean proportion of items correct after a time announcement for each of the groups of readers. This is essentially the null hypothesis, i.e., that the mean difference between scores before and after the time announcement is zero for each of the experimental groups.

It is possible, of course, that any change in performance might be due to factors other than the time announcement. One such factor might be increasing item difficulty as one goes through the test. Another factor which might tend to depress comprehension scores is fatigue. No direct measure of fatigue was obtained. Therefore, its influence can be considered only on a somewhat speculative basis in this study.

Rejection of the hypothesis at the 1% level of confidence for the slow readers but not for the fast readers may be interpreted as evidence in support of a position that the time announcement does have some effect on subsequent comprehension scores. The exploratory nature of the study justifies the setting of such a high level of confidence (**1**).

Procedure

Three groups of 100 answer sheets each were selected from approximately 1,500 answer sheets for the reading comprehension test administered in September 1955 using the percentile values assigned to the rate score as the criterion of selection.

The first group (Group A) consisted of examinees whose rate scores were above the 90th percentile and who had completed the test. The other two groups (B and C) consisted of slow readers. Group B was made up of examinees whose rate scores were below the 10th percentile and who had completed the test. Group C consisted of examinees whose rate scores were below the 15th percentile but who did not complete the test. The original intent was to select Group C from among the examinees in the lowest decile for the rate score, but the lack of a sufficient number of incompleted papers in this area necessitated selection from the area below the 15th percentile in order to secure 100 papers. It was assumed that the individuals in Group C, by not completing the test after the time announcement, either did not speed up or, if they did, they did so to a limited amount. This group was used as a sort of a "control" against which the performance of the slow readers

completing the test might be better judged. No corresponding group of unfinished papers was available for the faster readers since most of this group finished the test.

For each examinee in the three groups, the number of items correct before the time announcement was determined and expressed as a proportion of the items attempted in that period. A similar proportion was also determined for items correct of items attempted after the time announcement. A difference score was found for each examinee by subtracting the proportion correct before from the proportion correct after the time announcement. By this procedure, negative differences would indicate that the individual did not do as well (i.e., had a smaller proportion of items correct of items attempted) after the time announcement as before. The median number of items attempted before and after the time announcement was obtained for each group separately. The *t* test for related measures was used to test the significance of the mean difference for each group. No statistical tests of significance were made for differences between the groups.

Results

The results of the study are presented in Table 1. In this table are given the mean proportion of items correct of items attempted before the time announcement for each group, the mean proportion correct of items attempted after the time announcement, and the mean difference for each group. The median number of items attempted before and after the time announcement for each group is also presented in Table 1.

The results for the fast readers completing the test (Group A)

TABLE 1

ITEMS BEFORE AND AFTER TIME ANNOUNCEMENT

Group	Median Number Attempted		Mean Proportion Correct		
	Before Time	After time	Before time	After time	Diff.
A (*N* = 100)	63	17	.59	.55	−.04
B (*N* = 100)	37	43	.53	.39	−.14**
C (*N* = 100)	33	31	.58	.39	−.19**

** *t* significant at 1% level of confidence.

showed that they do read quickly as evidenced by the fact that over one half of the examinees had completed about 79% of the items before the time announcement. For this group, the mean proportions of items correct before and after the time announcement were quite similar. The mean proportion of items correct before the time announcement was .59. After the announcement, the mean proportion of items correct was .55. The average performance after the time announcement was in a negative direction but not statistically significant at the 1% level.

The results for the two groups of slow readers are quite similar. The slow readers completing the test (Group B) tended to speed up as predicted in that 53% of the items were attempted after, as compared to 47% before, the time announcement. The slow readers not completing the test (Group C) apparently were not greatly affected by the time announcement, as they completed 33 out of the 80 items before the time announcement but only 31 of the remaining 47 items after the time announcement. The slow readers completing the test had a mean proportion of correct answers of .53 before and .39 after the time announcement. Slow readers not completing the test had a mean proportion of .58 before and .39 after the time announcement. Both groups showed a quite similar "lack of success" after, as compared to before, the time announcement, the difference being less for slow readers completing the test than for slow readers not completing the test. The difference in mean proportional values before and after the time announcement were statistically significant at the 1% level.

Summary

The purpose of this study was to compare accuracy of comprehension scores secured for slow and fast readers before and after a time announcement used to secure a measure of reading rate. The specific hypotheses tested were that the mean difference between scores before and after the time announcement was zero for each of the experimental groups.

Fast readers completing the test had a mean difference of $-.04$. Slow readers completing the test had a mean difference of $-.14$, while slow readers not completing the test had a mean difference of $-.19$. Statistically significant differences between

the proportion of items correct of items attempted before and after a time announcement were obtained for groups of slow readers but not for fast readers.

These results suggest that a time announcement used to obtain a rate score is associated with more lowered comprehension scores after such an announcement than before it for slow readers but not fast readers, as measured by proportion of items correct of items attempted in the two respective time limits.

References

1. Anderson, Verna L., & Tinker, M. A. The speed factor in reading performance. *J. educ. Psychol.*, 1936, **27,** 621–624.
2. Flanagan, J. C. A study of the effect on comprehension of varying speeds of reading. In *Research on the Foundations of American Education.* Wash. D.C.: Amer. Educational Res. Ass., 1939. Pp. 47–50.
3. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton-Mifflin, 1953.
4. Preston, R. C., & Botel, M. Reading comprehension tested under timed and untimed conditions. *Sch. & Soc.*, 1951, **74,** 71.
5. Robinson, F. P., & McCollum, F. H. Reading rate and comprehension accuracy as determinants of reading test scores. *J. educ. Psychol.*, 1934, **25,** 154–157.
6. Stroud, J. B. A critical note on reading. *Psychol. Bull.*, 1942, **37,** 173–178.
7. Stroud, J. B., & Henderson, Margaret. Rate of reading and learning by reading. *J. educ. Psychol.*, 1943, **34,** 193–205.
8. Tinker, M. A. Speed versus comprehension in reading as affected by level of difficulty. *J. educ. Psychol.*, 1939, **30,** 81–94.
9. Tinker, M. A. The relation of speed of comprehension in reading. *Sch. & Soc.*, 1932, **36,** 1958–160.
10. Traxler, A. E. *Ten years of research in reading,* Bull. 32. New York: Educational Records Bureau, 1941.
11. Traxler, A. E., & Townsend, Agatha. *Another five years of research in reading,* Bull. 46. New York: Educational Records Bureau, 1946.
12. Traxler, A. E., & Townsend, Agatha. *Eight more years of research in reading,* Bull. 64. New York: Educational Records Bureau, 1955.

Received May 15, 1957.

UNDER- AND OVER-AGE GRADE-PLACEMENTS AND SCHOOL ACHIEVEMENT

JACK A. HOLMES

University of California, Berkeley

AND CARMEN J. FINLEY

Sonoma School Districts, California

This study concludes a series of investigations at the elementary school level designed to establish relationships between the degree of over- and under-age grade-placement and relative success in various school subjects. The first two studies, on Grades V and VI, have already been reported (**3, 4**). The present paper presents data for Grades VII and VIII and then discusses the theoretical significance of the findings.

The hypothesis being tested holds that certain academic disciplines at the elementary level are more important than others in contributing to nonpromotion, promotion, and acceleration.

Samples, Methodology, Tests

For Grade VII, 293 boys and 312 girls and for Grade VIII, 318 boys and 293 girls were drawn from pupils enrolled in 68 elementary school districts in Sonoma County.

Grade-Placement Deviation (GPD) was determined by computing the difference between each child's Actual Grade-Placement (AGP) and the grade he should have been in according to his chronological age. This latter grade is defined as his Chronological-age Grade-Placement (CaGP). Grade-Placement Deviation, then, is expressed as:

$$GPD = AGP - CaGP + K$$

where K is a constant (5.0) inserted into the equation to give all GPD values a positive sign.

Table 1 gives the frequency distributions of GPD's for each sex in both grades. Inspection reveals that, according to their

TABLE 1

FREQUENCY DISTRIBUTIONS OF ACTUAL GRADE-PLACEMENT DEVIATIONS

Direction of Deviation	GPD Interval	Grade VII		Grade VIII	
		Boys f	Girls f	Boys f	Girls f
(Under-aged)	2.3 + 2.8[a]	0	1[b]	0	0
(Accelerated)	1.7 + 2.2	0	0	0	0
↑	1.1 + 1.6	4	9	3	6
↑	.5 + 1.0	49	72	43	70
Normal	− .1 + .4	125	140	124	123
↓	− .7 − .2	70	59	83	55
↓	−1.3 − .8	28	16	42	25
↓	−1.9 − 1.4	12	11	20	8
↓	−2.5 − 2.0	3	3	3	4
(Over-aged)	−3.1 − 2.6	1	1	0	2
(Retarded)	−3.7 − 3.2	1	0	0	0
N		293	312	318	293

[a] The constant K has been subtracted in order to bring these deviations from the mean back into their original form.

[b] Example: This seventh grade girl is accelerated by +2.3 grades; therefore, according to her chronological age she should have been in the (7.1 − 2.3) or 4.8th grade.

chronological ages, the seventh-grade boys *should* have been distributed from Grades 5.7 to 10.4 and the seventh-grade girls should have ranged from 4.8 to 9.7; yet, the entire group was actually in grade 7.1.

Likewise, Table 1 shows that, in accordance with their chronological ages, the boys and girls of the eighth grade should have been enrolled in Grades 6.5 to 10.6 and 6.5 to 11.2, respectively; yet, they were all in Grade 8.1.

The California Achievement Test Battery, Intermediate Form DD, was administered to all seventh- and eighth-grade children during the first three months of the school year as part of the general county testing program in 1952. The tests in this battery are: reading vocabulary, reading comprehension, arithmetic reasoning, arithmetic fundamentals, mechanics of grammar, and spelling. Reliability coefficients range from .85 to .95.

TABLE 2

INTERCORRELATIONS OBTAINED ON 293 BOYS IN GRADE VII

Variable	1 GPD	2 RV	3 RC	4 AR	5 AF	6 MG	7 SP
1. Grade Placement Deviation (GPD)		.450	.346	.318	.259	.364	.397
2. Reading Vocabulary (RV)			.750	.650	.533	.622	.704
3. Reading Comprehension (RC)				.637	.585	.655	.670
4. Arithmetic Reasoning (AR)					.667	.569	.538
5. Arithmetic Fundamentals (AF)						.564	.496
6. Mechanics of Grammar (MG)							.540
7. Spelling (SP)							
Mean	6.98[a]	6.98	6.75	6.78	6.15	6.70	6.83
Standard Deviation	.68	1.60	1.39	1.01	1.04	1.30	1.65

[a] Reconverted: K-Constant Subtracted.

Findings

Grade VII

Table 2 presents the intercorrelations found among the above variables and Grade-Placement Deviation for the boys. The criterion correlates highest with reading vocabulary ($r = .450$) and lowest with arithmetic fundamentals ($r = .259$).

The matrix in Table 2 was submitted to the Wherry-Doolittle Test Selection Method (7) in order to determine which of the test variables made an independent contribution to the variance of the pupils' GPD's. The multiple correlation ($\bar{R}$) rises as the variables, reading vocabulary, spelling, and mechanics of grammar, are successively selected and cumulatively added, but then decreases if any of the other tests are added. Together these three tests yield a multiple $\bar{R}$ of .46, significant at better than the 1% level of confidence.

Table 3 presents the intercorrelations for the girls. The criterion (GPD) correlates highest with reading vocabulary ($r = .380$) and lowest with arithmetic fundamentals ($r = .233$). A Wherry-Doolittle analysis selects only reading vocabulary and spelling

TABLE 3

INTERCORRELATIONS OBTAINED ON 312 GIRLS IN GRADE VII

Variable	1 GPD	2 RV	3 RC	4 AR	5 AF	6 MG	7 SP
1. Grade Placement Deviation (GPD)		.380	.355	.278	.233	.269	.332
2. Reading Vocabulary (RV)			.770	.569	.477	.644	.619
3. Reading Comprehension (RC)				.636	.578	.703	.670
4. Arithmetic Reasoning (AR)					.618	.574	.489
5. Arithmetic Fundamentals (AF)						.572	.492
6. Mechanics of Grammar (MG)							.598
7. Spelling (SP)							
Mean	7.13[a]	7.18	7.23	6.75	6.40	7.30	7.62
Standard Deviation	.68	1.40	1.34	.84	1.16	1.27	1.68

[a] Reconverted: K-Constant Subtracted.

TABLE 4

INTERCORRELATIONS OBTAINED ON 318 BOYS IN GRADE VIII

Variables	Intercorrelations						
	1 GPD	2 RV	3 RC	4 AR	5 AF	6 MG	7 SP
1. Grade Placement Deviation		.434	.440	.430	.327	.381	.439
2. Reading Vocabulary (RV)			.781	.691	.564	.667	.699
3. Reading Comprehension (RC)				.723	.617	.651	.619
4. Arithmetic Reasoning (AR)					.751	.612	.549
5. Arithmetic Fundamentals (AF)						.634	.517
6. Mechanics of Grammar (MG)							.553
7. Spelling (SP)							
Mean	7.88[a]	7.78	7.51	7.75	7.05	7.40	7.58
Standard Deviation	.68	1.56	1.57	1.31	1.46	1.38	1.74

[a] Reconverted: K-Constant subtracted, remainder algebraicly added to 8.1.

TABLE 5

INTERCORRELATIONS OBTAINED ON 293 GIRLS IN GRADE VIII

Variables	Intercorrelations						
	1 GPD	2 RV	3 RC	4 AR	5 AF	6 MG	7 SP
1. Grade Placement Deviation		.348	.365	.302	.298	.279	.327
2. Reading Vocabulary (RV)			.774	.641	.571	.652	.634
3. Reading Comprehension (RC)				.688	.647	.654	.655
4. Arithmetic Reasoning (AR)					.786	.665	.580
5. Arithmetic Fundamentals (AF)						.656	.572
6. Mechanics of Grammar (MG)							.600
7. Spelling (SP)							
Mean	8.11[a]	8.18	8.03	7.85	7.65	8.14	8.67
Standard Deviation	.71	1.49	1.41	1.28	1.62	1.42	1.67

[a] Reconverted: K-constant subtracted, remainder algebraicly added to 8.1.

as variables making independent contributions to the variance of GPD's for seventh-grade girls. Together, the two tests yield an $\bar{R}$ of .393, significant beyond the 1% level.

Grade VIII

Table 4 presents the intercorrelations for the boys. GPD correlates highest with reading comprehension ($r = .440$) and lowest with arithmetic fundamentals ($r = .327$). The multivariate analysis selects reading comprehension, spelling, and arithmetic reasoning. Together, these yield an $\bar{R}$ of .50, significant beyond the 1% level of confidence.

Table 5 presents the intercorrelations for the girls. GPD correlates highest with reading comprehension ($r = .365$) and lowest with mechanics of grammar ($r = .279$). Analysis (7) selects reading comprehension, spelling, and reading vocabulary. Together, these yield an $\bar{R}$ of .38, significant at the 1% level.

Discussion

In order to compare the results obtained in Grades VII and VIII with those found for Grades V and VI already published

TABLE 6

SUMMARY OF SELECTED TESTS AND CUMULATING $\bar{R}$'S FOR GRADES V, VI, VII, AND VIII

Grade	Boys		Girls	
	Tests	Cumulative $\bar{R}$	Tests	Cumulative $\bar{R}$
V	Reading Vocabulary	.248	Spelling	.170
	Spelling	.268[a]	Reading Vocabulary	.183
	Mechanics of Grammar	.275[b]	Arithmetic Fundamentals	.190
VI	Spelling	.365	Reading Vocabulary	.302
	Arithmetic Reasoning	.393	Spelling	.321
	Mechanics of Grammar	.394	Reading Comprehension	.326
VII	Reading Vocabulary	.447	Reading Vocabulary	.377
	Spelling	.458	Spelling	.393
	Mechanics of Grammar	.464		
VIII	Reading Comprehension	.437	Reading Comprehension	.361
	Spelling	.483	Spelling	.375
	Arithmetic Reasoning	.497	Reading Vocabulary	.379

[a] Cumulative $\bar{R}$ for Reading Vocabulary + Spelling.

[b] Cumulative $\bar{R}$ for Reading Vocabulary + Spelling + Mechanics of Grammar.

(3, 4), a summary of the selected tests and the cumulative multiple correlations found for both sexes at all four levels is presented in Table 6. Inspection of Table 6 reveals that:

1. Spelling is selected as an important test making a statistically independent contribution to the variance of GPD in all four grades for both boys and girls. The same cannot be said for any other variable in the present battery.

2. Reading vocabulary and spelling are the two most important determiners of GPDs in Grades V, VI, and VII.

3. The cumulative $\bar{R}$ tends to increase from grade to grade.

4. The cumulative $\bar{R}$ is greater for boys than for girls at each grade level.

TABLE 7

RANK-ORDERS FOR SCHOOL SUBJECTS PLAYING MOST TO LEAST IMPORTANT ROLES IN DETERMINING GRADE-PLACEMENT DEVIATIONS

Calculated from College Students' Estimates	Estimated from the $\bar{R}$ Calculations
Grades V, VI, VII	
(N = 75)	(Ns = 979 + 869 + 611)
1. Reading Comprehension	1. Reading Vocabulary
2. Reading Vocabulary	2. Spelling
3. Arithmetic Fundamentals	3. Mechanics of Grammar
4. Arithmetic Reasoning	4. Reading Comprehension
5. Spelling	5. Arithmetic Reasoning
6. Mechanics of Grammar	6. Arithmetic Fundamentals
Grade VIII	
(N = 75)	(N = 611)
1. Reading Comprehension	1. Reading Comprehension
2. Arithmetic Reasoning	2. Spelling
3. Reading Vocabulary	3.5 Arithmetic Reasoning
4. Mechanics of Grammar	3.5 Reading Vocabulary
5. Arithmetic Fundamentals	5. Arithmetic Fundamentals
6. Spelling	6. Mechanics of Grammar

5. Mechanics of grammar is selected as an important subject for boys in grades V, VI, and VII, but this is not so for the girls.

6. Individual differences in achievement in the *quantitative* subjects (arithmetic reasoning and arithmetic fundamentals, in which boys are supposed to excel) appear to play almost no role in determining whether or not a student shall be retarded, promoted, or accelerated in the elementary schools of this system!

7. Individual differences in achievement in the *linguistic* sub-

jects (reading, spelling, and grammar, in which girls are supposed to excel) have a high premium in regard to GPD.

Could these results have been logically predicted?

Perhaps. The fact is, however, they were not! A class of 75 university students enrolled in a teacher-training program was asked to rate these subjects in the order of their importance for determining GPD's in Grades V through VIII.

Table 7 presents the rank orders obtained by preferential analysis of the estimates of these university students and compares them with those obtained from a preferential analysis of the $\bar{R}$'s of the four grades. The university students stress comprehension and relegate spelling to a low order of importance. Their disagreement with the facts is readily apparent in the almost random relationship of the rank orders of the entries in the two columns, especially for Grades V, VI, and VII.

Theoretical Implications

In essence, these studies recall Keys' observations (**5, 6**) that acceleration is rare; indeed, the number of pupils that are accelerated in any one grade represent only a fraction of those retarded. The writers agree with Keys' contention that this does not make for the optimum learning situation for the bright, the average, or the dull. The usual argument for the policy of non-acceleration is that the child should be kept with his chronological-age group. Such a social milieu is supposed to foster both maximum mental health and learning. However, these data clearly show that the desired homogeneity for chronological age (or for that matter, subject-matter achievement) is simply not obtained.

Baker (**1**) and Baker and Traphagen (**2**) concluded from their studies that often bright pupils may fail in arithmetic, and spelling failure may occur at any level. The relationship of their findings to the present study is clear: bright or dull failure in arithmetic is not important for GPD; and spelling failure, which they found occurs at any ability level, also occurs at any grade level. Its occurrence is of great importance for GPD.

The theoretical implications of the present series of investigations concerning school-grade failures seem clear. The schools should come out openly and tell the students and their parents

that success in reading comprehension and arithmetic reasoning is not nearly as important as success in reading vocabulary, spelling, and mechanics of grammar for determining retardation, promotion, or acceleration by the time the child is in Grades V, VI, and VII. *Or*, the schools should make it their business to give a more equable distribution to the weights they place on success in each of the subject-matter fields.

As a matter of fact, for neither the boys nor the girls, in Grades V, VI, VII, and VIII, do individual differences in their *combined achievement* in these six subject-matter areas (i.e., reading vocabulary, spelling, mechanics of grammar, reading comprehension, arithmetic reasoning, and arithmetic fundamentals) contribute more than 25% to the variance of the grade-placement deviations within any one class. On what basis, then, one may ask, are students being grouped together in any one grade? On what basis are they being promoted, retarded, or accelerated?

In a scientific age, the apparent unimportance of success in arithmetic is perplexing. This condition is no doubt widespread in the nation's schools, and it is quite possible that this is one of the major reasons why officials of the armed services complain that a great number of men are not well grounded in arithmetic.

If these data are, in fact, at all representative of elementary schools in general, they point up a pressing need for a re-evaluation of the relative importance of the courses offered in the school curriculum.

In an age of interplanetary exploration, the outcome of the race for world leadership may not depend so much upon the ability of a nation to successfully project a satellite into a precalculated orbit, nor even its ability to launch intercontinental missiles, but, as Professor Edward Teller has so clearly stated, upon its ability to simply take and hold the lead in scientific exploration and production. If the results of the above experiment are at all widespread, and there is reason to believe that they are, then it appears that there must be a revamping of our value judgments on what in the curriculum is of the most worth. More than that, knowledge for its own sake, especially in mathematics and science, must be much more enthusiastically supported and *evaluated* as important accomplishments for children in our elementary schools.

References

1. Baker, H. J. *Characteristic differences in bright and dull pupils.* Bloomington, Ill.: Public Sch. Publishing Co., 1927.
2. Baker, H. J., & Traphegen, V. *The diagnosis and treatment of behavior problem children.* New York: Macmillan, 1935.
3. Holmes, J. A., & Finley, Carmen J. Relative importance of curricular areas for grade-placement deviation in grade V. *Calif. J. educ. Res.*, 1955, **6**, 213–218.
4. Finley, Carmen J., & Holmes, J. A. Relative importance of curricular areas for grade-placement deviation in grade VI. *Calif. J. educ. Res.*, 1956, **7**, 200–205.
5. Keys, N. The underage student in high school and college. *Univer. Calif. Publ. Educ.*, 1938, **7**, 145–272.
6. Keys, N. Should we accelerate the bright? *J. except. Child.*, 1942, **8**, 248–254.
7. Wherry, R. J. A new formula for predicting the shrinkage of co-efficient of multiple correlation. *Ann. math. Statist.*, 1931, **2**, 440–457.

Received June 5, 1957.

ITEM DIFFICULTIES BASED ON END SEGMENTS

EDWARD L. CLARK

Northwestern University

In the analysis of test items, there is a strong tradition, originated by Kelley (**3**), for using the proportion of right answers in the top and in the bottom 27% of *S*s when ordered by a criterion. These end groups are used to determine the index of discrimination of the items, and some writers recommend the use of these end groups, neglecting the middle 46%, for determining the difficulty of the items (**1**). This use of the top and the bottom "quarters" of the criterion group for determining the difficulty of test item fails to distinguish between those items which can be answered by only the top 27% and those items which can be answered correctly by as many as the top 73%. Feder (**2**), using the responses of 144 *S*s to the 70 items of a classroom test, has offered evidence in support of determining item difficulty by the end "quarters" of the criterion group. He concluded not only that the difficulty could be determined by the end groups, but also that the usefulness of this method was not related to the discrimination index of the item.

The purpose of the present study was to compare, by correlation, the indices of difficulty as determined by the end quarters of the criterion group with those as determined by the middle two quarters. In order that the effectiveness of the items might be a variable in the study, correlations between difficulties were made for items of five different levels of discrimination.

The data used in this investigation were 1911 item analyses which had each been based on 400 *S*s. These records had been accumulated over a 12-year period as two scholastic aptitude tests were being constructed: the Clark Analogies Test, in parallel forms and based on a very easy vocabulary, and a second test based on a difficult vocabulary. All items were of the multiple-choice type. The criterion most often used in these original analyses was total test scores. At times it was college grades or a mixed criterion of test scores and grades. The poorest items were

TABLE 1

DIFFICULTIES AS DETERMINED BY THE TWO END QUARTERS AND BY THE MIDDLE HALF OF GROUPS

Phi Coef.	No. of item analyses	End groups Mean	End groups *S D*	Middle groups Mean	Middle groups *SD*	*r*	*SD* of all
0–.24	278	115.0	51.94	114.0	55.89	.984	107.38
.25–.39	816	124.0	36.68	122.8	48.83	.977	85.03
.40–.49	497	118.6	29.26	120.9	39.07	.956	67.59
.50–.59	245	113.6	21.15	116.4	22.20	.912	42.42
.60–up	75	107.7	16.30	105.1	27.40	.872	42.38

Note.—The means are the mean number of correct responses to items by 200 *S*s, while the correlation coefficients are between the two determinations of difficulty for each item.

generally analyzed but once; the better items were analyzed with several groups of *S*s. Although five different indices of discrimination were used at different times, the only index used throughout the period was the phi coefficient based on the number of right answers made by the highest 100 and the lowest 100 *S*s. Consequently, for the purposes of this study the phi coefficient was used to classify the items as to level of discrimination.

The first line of data of Table 1 concerns those items which yielded phi coefficients of less than .25, or items which did not discriminate between the highest and lowest *S*s of the criterion groups. In these 278 analyses of items, the end groups of criterion *S*s made right answers in almost the same manner as did the middle groups. The mean and standard deviation of right answers per item were much the same for each group of 200 *S*s, and the correlation between them was $r = .984$. The standard deviation of the number of right answers made by the entire group was 107.4. These items varied greatly in difficulty.

For items of all levels of discrimination, the average number of correct responses is somewhat above 100, out of a possible 200, but for the items of high discrimination level the standard deviations are much lower. Items which discriminate well between strong and weak *S*s not only have an average difficulty near 50% but individualy are much closer to this level of difficulty.

Some importance should be attached to the fact that the standard deviation of the number right per item for the total group of 400 is more than twice as great as the corresponding standard

deviations based on the answers of the end groups. The greater relative variability of item difficulties suggests strongly that more information is being given by the entire group than by the ends alone. Since the standard deviations for the middle group were consistently larger than for the ends, the middle half of *Ss* gave better evaluations of difficulty for a majority of the items. However, of more concern in this study are the several coefficients of correlation between the difficulties as determined by the end groups and the middle group. While these correlations are very high, they get progressively lower with the increasing discrimination level of the items. In passing, it should be noted that the use of 400 *Ss* in each analysis gave rather stable halves and allowed the highest correlations to approach 1.00. The use of 800 *Ss* would doubtless have given even higher correlations and the use of only 100 would certainly have given lower. But the important finding in this study is that the use of the end "quarters" alone for determining item difficulties becomes progressively less satisfactory as one deals with increasingly better items.

The use of the proportion of right answers in the end segments of the criterion group for determining the index of discrimination is admittedly an approximation method, but this fact, in this age of card-sorting machinery, seems little justification for not using all the data at hand for determining item difficulties. However, if large numbers of *Ss* are used and *if one's items are of little discrimination value,* the item difficulties may quite accurately be determined from the top and the bottom "quarters" of the criterion group.

References

1. Davis, F. B. Item-analysis data; their computation, interpretation, and use in test construction. *Harv. Educ. Pap.,* 1946, No. 2.
2. Feder, D. D. Estimating of test item difficulty by averaging highest and lowest quarter performance compared with total population count. *Educ. psychol. Measmt,* 1947, **7,** 133–134.
3. Kelley, T. L. The selection of upper and lower groups for the validation of test items. *J. educ. Psychol.,* 1939, **30,** 17–24.

Received October 9, 1957.

THE INTERRELATIONSHIP OF AREA TEST SCORES AND COOPERATIVE GENERAL CULTURE TEST SCORES

JOHN C. WRIGHT

DePauw University

AND BARRON B. SCARBOROUGH

Florida State University

In the fall of 1954, Educational Testing Service (ETS) made available the Area Tests. These tests were "... designed to assess the broad outcomes of education in the liberal arts" (**4**). A description of the tests further stated that "... the tests have been aimed beyond the details of specific courses in an attempt to reach toward those goals of education which are shared generally by the colleges offering instruction in the liberal arts" (**3**). The Cooperative General Culture Test, also available from ETS, was designed "... to provide an indication of the student's general background in five important areas of the college curriculum" (**4**).

In February 1955, the Area Tests were administered at DePauw University in an effort to gain an over-all estimate of student progress in a liberal arts setting. These tests were administered since an estimate was desired of student progress in acquiring a way of thinking and dealing with problems which transcended the mere recall of a volume of relatively isolated academic information. The Cooperative General Culture Test was not used since there was some evidence to indicate it merely attempts to assess the student's knowledge of a wide range of academic, factual information (**2**). The Cooperative General Culture Test is not considered an adequate measure of liberal education (**1**).

Inasmuch as the Cooperative General Culture Test had been administered two years previously to the same group of students, there was a question of the distinctiveness of the two instruments. Therefore, this study was undertaken to see what relationship existed between the scores of students on the Area Tests and the scores of students on the Cooperative General Culture Test at

DePauw University. Seemingly, the two tests were designed for quite different purposes.

Procedure

In February 1955, the Area Tests (Form I) were administered to the senior class of DePauw University ($N = 269$). These tests were returned to ETS for scoring. Scaled scores were furnished for each student. Of the 269 seniors who took the Area Test, 231 were administered the Cooperative General Culture Test in February 1953.

Conditions of administration were comparable in that the same room was used, both tests were part of a university testing program, and the test administrator in each instance had had years of group testing experience.

Product-moment correlations were computed for the scores of each of the subtests of the Cooperative General Test and the related area of the Area Tests. Thus, the scores for Current Social Problems and History and Social Studies were correlated with the score for the area of Social Studies. The scores for Literature and Fine Arts were correlated with the score for the area of Humanities. The scores for Science and Mathematics were correlated with the score for the area of Natural Science. The intercorrelations were obtained for each sex and for the sexes combined. In all instances, raw scores were used for the Cooperative General Culture Test scores and scaled scores used for the Area Tests scores. Raw scores for the Area Tests were not available.

Results

The intercorrelation indices are shown in Table 1.

In general, the correlation coefficients were higher for the women than for the men. None of the coefficients for the women was below .60 while two of these were in the seventies. The lowest single coefficient was .53. This index, which was for the men, was between Mathematics and the area of Natural Science. Both of the coefficients for the area of Humanities for men were .58. All other coefficients for both men and women were in the sixties.

Discussion

The obtained coefficients of interrelationship of the Area Tests and the Cooperative General Culture Tests were, in general, rela-

TABLE 1

CORRELATIONS OF AREA TEST SCORES AND GENERAL CULTURE TEST SCORES

General Culture Test	Area Test	Senior Men (N = 123)	Senior Women (N = 108)
Current Social Problems	Social Studies	.64	.66
History & Social Studies	Social Studies	.64	.75
Literature	Humanities	.58	.61
Fine Arts	Humanities	.58	.72
Science	Natural Science	.66	.66
Mathematics	Natural Science	.53	.60

tively high. The coefficients were not as high as might be expected for a two-year interval on a test-retest basis for either test.

Many possibilities exist which might account for the obtained coefficients, including: the acquisition of certain facts is a prerequisite to the utilization of other facts, certain factors are not measured but common to each test used (e.g., general intelligence), and/or there is a communality of factors measured. It is possible that DePauw University students constitute an atypical liberal arts sample. Whatever the reason, the evidence suggests further investigation of the distinctiveness of the two instruments.

Summary

The Area Tests and the Cooperative General Culture Tests were designed for somewhat different purposes. The former was designed to assess the broad outcomes of a liberal arts education, whereas the latter was designed to estimate the general background of the student in certain areas of the college curriculum. This study is an estimate of the interrelationship of the subtest scores of the Area Tests and the scores of the Cooperative General Culture Test for DePauw University students. In February 1955, the Area Tests were given to the senior class of DePauw University. The Cooperative General Culture Test had also been administered to the senior class as sophomores in February 1953. Intercorrelations were computed between the comparable areas of each instrument. The obtained coefficients were relatively high. A number

of possibilities might account for the indices. However, the evidence seems sufficient to warrant further investigations of the distinctiveness of the two tests.

References

1. Buros, O. K., (Ed.) *The third mental measurement yearbook,* New Brunswick, Rutgers Univer. Press, 1949.
2. Buros, O. K., (Ed.) *The fourth mental measurement yearbook,* New Jersey: Gryphon Press, 1953.
3. Educational Testing Service *Institutional testing program; handbook for deans and examiners.* Princeton, N. J.: Author, 1954–55.
4. Educational Testing Service. *Catalog of cooperative tests,* Princeton, N. J.: Author, 1956.

Received October 9, 1957.

Book Reviews

DANIEL BROWER, & LAWRENCE E. ABT. *Progress in Clinical Psychology*. New York: Grune & Stratton, 1956. Pp. 364.

This volume fulfills the promise made in the first, which presented progress in clinical psychology through 1951. This study includes primarily the years 1952, 1953, and 1954, with several subjects not appearing in the first volume going back as far as 1946.

The introduction deals with clinical psychology from a transactional point of view; Part 2 includes discussions of the Rorschach method, thematic apperception and fantasy tests, and validity of sentence completion tests and human figure drawings. Part 3 discusses therapy in relation to psychoanalysis, client-centered approach, group methods, reading and other subject disabilities, speech disorders, play, music, art, bibliotherapy, psychodrama, hypnosis, and dreams. Part 4 reviews clinical psychology in correctional institutions, rehabilitation of the physically handicapped, and administrative problems in state institutions. Part 5 takes up problems of neuropsychology, statistics and research.

The new subjects not included in Volume 1 are hypnosis, dreams, speech, reading, and some specialized therapeutic techniques.

The very condensed nature of the studies may be seen from the fact that one sentence may have from three to six references, and eight pages of text may have six pages of bibliography. Many statements are references to other summaries. Including all the bibliographies at the ends of chapters, there are 1560 references. The ten subjects with the most extensive bibliographies are the following: dreams (178 references); group psychotherapy (167); rehabilitation of physically handicapped (163); statistics, reading, etc. (each 155); neuropsychology (114); psychoanalytic techniques (75); play, etc. (70); research (64); hypnosis (60). These numbers must not be taken as a measure of anything, but they perhaps suggest the interest of workers in the past few years as represented by the many authors of the present volume—something which might have been considerably different if there had been other writers.

While giving considerable amounts of information, the study primarily refers the advanced student to most of the best in the field. This is not for beginning students, but implies a foundation in the subject. Written clearly in general, it shows some of the

ambiguity, paradoxes, and contradictions now existing. Even with these writers we do not quite get away from "faculties" and "fixed tools of thought." Some of the ambiguity in the field of clinical psychology seems, in the language of the psychology of learning, to have transferred.

In general these are scholarly reports by specialists writing in their separate fields of interest and bring the reader up to date (through 1954) in clinical psychology, probably as carefully and accurately as the difficulties of the subject matter permit.

Of special merit are the wide scope of bibliographical material, excellent critical comments, and suggestions for improvement in procedure and for further research, warnings against faulty and unscientific methods, and indications of the gaps that need more attention and filling in.

A good index of subjects completes the volume.

A. S. Edwards

The University of Georgia

Frieda Fromm-Reichmann, & J. L. Moreno (Eds.) *Progress in Psychotherapy*, 1956. Grune & Stratton, 1956. Pp. 352.

This is a presentation by about 45 writers and collaborators of the philosophies, work, and progress of the various schools of psychotherapy. It grew out of the "program offered by our new Section at the annual meeting of the American Psychiatric Association in Atlantic City, 1955." Additions were made to include schools which were not heard at the meeting. It is hoped that it will help "to offer a platform for unprejudiced oral and written presentation and comparison of data by representatives of all schools of psychotherapeutic thinking."

Part 1 presents something of the history and philosophy of psychotherapy. Part 2 deals with principles of psychotherapy. Part 3 presents statements by workers in psychotherapy who indicate the nature and methods each of his own school. Part 4 gives summaries of developments in seven foreign countries. Part 5 is a summary of the present status and possible future of psychotherapy.

Part 3, dealing with various schools, includes discussions on the rediscovery of the patient, Adlerian psychotherapy, psycho-

biologic psychiatry, Jung's concepts, active analytic therapy, existential analysis, psychodrama, the work of Sullivan and Fromm, the views of Horney, communication, biodynamic integration, client centered therapy, multiple therapy, hypnotherapy, countertransference, problems of schizophrenia, group psychotherapy, and prevention and treatment for psychotic children.

An immense number of techniques are indicated; in fact, in his summary, Moreno states that "One hundred and forty-two techniques are mentioned in this book." They may, according to the same author, be classified under seven headings: "physical, pharmacologic, hypnotic, individual, group, action, and technological psychotherapy." The immense sweep and inclusiveness of the subject matter is matched by general excellence of most of the articles. This might be expected since each school in the major presentations is discussed by a selected representative and worker in that particular school. For the foreign countries, Austria, Great Britain, France, Germany, South America, Spain, and Switzerland, the summaries are generally written by one person.

In general, there appear to be very frank and humble statements with clear recognition, at least by many, of how little we know and how far we have to go in understanding psychotherapy. It may appear to many readers that the book contains an amazing collection of unverified concepts and reified hypotheses. Many clear distinctions are made, but we can even find suggestion that psychotherapy and psychoanalysis are indistinguishable. There are many statements of problems, successes and failures; but there are also suggestions of arrogance existing in the field, of unacceptable dogmatisms, and broad generalizations on insufficient data. In the multiplicity of attempts to find foundations for psychotherapy, it is refreshing to find Masserman referring to sound experimental work. In some cases it would seem that experimental psychology and general psychology of the textbooks were treated with thorough disregard if not with contempt. Some formulations seem to be straight from the armchair.

It is pointed out that practice is probably better than theory and definition, and that much of successful therapy lies in the personality and competence of the therapist. The heavier weighting of the volume is on psychoanalytic concepts; rather little recognition is given to interpretation in terms of learning. The experimental psychologist reading this book might like to ask how

far astray can we get when imagined entities are substituted for experimental facts? He would not be favorably inclined towards "omnipotent (unconscious) elements;" and he would be inclined to agree that there is too much jargon, too many "high-falutin words," and the danger of too many "parochial attitudes."

Some of these suggestions of adverse criticism have been frankly recognized by some of the writers and properly deprecated. Lack of scientific procedure is recognized and admitted in some of the presentations; but many of the statements are very carefully worked out with due consideration for scientific psychology and for not letting theory and philosophy of psychotherapy get too far out of hand.

It would be difficult to find any book which would better put the reader in touch with the psychotherapy of 1956, or, should we say, with the psychotherapies of 1956. Two indexes are well made and put the reader in touch with most of the details of the book.

A. S. Edwards

The University of Georgia

J. E. Wallace Wallin. *The Odyssey of a Psychologist.* Lyndalia, Delaware: Author. 1955. Pp. 243.

The author, who has more or less restrained himself in more than three hundred publications (books, journal articles, reviews and news articles), now presents what he wants to say about the efforts of half a century in the interest of handicapped children. The book is dedicated to the very many who "rendered the Author Unflagging Support and Encouragement In the Implementation of Services for Children Handicapped in the Race for Life by Accident, Disease, or Inheritance."

It is an autobiography and has two special concerns: first, "to set forth in some detail the nature of the factors that were influential in the making of a special educationist and clinical psychologist—in fact, that proved to be indispensable for the accomplishment of this purpose; second, to provide some insight into the intricate and complex problems of personal and professional adjustment in the field of education that may retard the progress of the work, or even derail it—not because of any flaw in the work itself but because of human weakness—and lead to waste of public funds and blighted hopes and disappointments."

The report includes personal experience in and knowledge about 31 training schools, colleges, and universities in 18 states. It covers also most of the period of the development of the work to which he dedicated his life and to which he probably gave more of pioneering efforts than has any other person. It is the story of his "pioneering experiences in special education, clinical psychology, and mental hygiene."

More or less chronologically, Wallin recounts his experiences and observations in schools and colleges; successes and failures; sources of encouragement and of discouragement; obstacles, both personal and impersonal; the defense of vested rights; progress and collapse of programs; difficulties in getting bills passed in legislatures; the folly of some psychologists; administrative fairness and unfairness; loyal support and strong opposition; untrained psychologists; personal views, both favorable and unfavorable of individuals; attitudes of M.D.s and the improving relations between psychiatrists and clinical psychologists; some of the ills of schools, of extension work, and of publication.

The Odyssey is interesting as a story; it is significant as giving much of the history of the development of clinical psychology and of special classes for handicapped children. Many details found in this volume have probably never been printed before.

The desire of Wallin was not only to give accurate information, but to render assistance to the curing of ills that still continue: "Little hope exists of removing unwholesome conditions that hamper the efficient functioning of public institutions unless the conditions are frankly revealed to the professions concerned and to the public." And he adds: "Fortunately, much in the recital redounds to the credit, honor, and glory of those who have contributed of their might to pioneering in the specific areas in which I have labored."

This is a significant and interesting supplement to the history of psychology, especially in connection with its applications to human improvement. It appears at a time when the work to which he dedicated his life is receiving greatly increased interest, effort, and substantial support.

The book includes almost a complete bibliography of Wallin's publications (Pp. 217–243).

A. S. Edwards

The University of Georgia

GREGORY A. KIMBLE. *Principles of General Psychology.* New York: Ronald Press, 1956. Pp. 400.

This appears to be a very sincere and conscientious effort to present beginning psychology, with great emphasis upon behavior and the extremely objective point of view. Psychology is defined as the "science of behavior." It deals only with "the aspects of behavior that are observable." Experience, mind, and consciousness are excluded from the definition, but introspective reports are acceptable. "It is not the subject's private experience which counts. It is, rather, the report of such experiences which is open to public inspection, verifiable, and testable. Reports of experience made by introspective psychologists are quite as acceptable as any other scientific records."

The presentation is in four parts: 1. Basic Methodology, including general orientation, psychology as a science, graphic and statistical methods, intelligence and aptitude testing, and functions of intelligence; 2. Sensation and Perception: sensory processes in general, vision and audition, and perception; 3. The Modification of Behavior: psychological development, conditioning and simple learning, selective learning, verbal and motor learning, extinction, habit breaking, and forgetting; 4. Behavior dynamics: the primary drives, secondary motivation, frustration, stress and conflict, personality and adjustment, and behavior pathology.

The author at the outset attempts to clarify the meaning and work of psychology as a science and to correct common misconceptions about psychology. The book proceeds to deal with psychology strictly as an objective science, emphasizing the S-R and R-R formulae but not S-O-R. Emphasis is placed upon understanding the place of theory, and of experiment, and of the refinement of experimental control in real science.

There is an appropriate exposition of graphical methods and a brief statement of statistics in which the median is unfortunately defined as "the middle score in a distribution."

The presentation of psychological facts and principles pays thorough respect to experimental work and admits the S-P-R paradigm in connection with perception. At least the word "experience" leaks in for, "Perception refers to more complex experience in which factors such as learning and motivation often play an important role." Of special note are the 211 well-selected illustrations

which not only help in clarification but add much of interest to the book, which in general is simply and clearly written.

The student will be interested in the discussions of behavior dynamics and of mental disorders. The author has attempted to give a careful evaluation of Freud's psychoanalysis. But it will be hard for many to classify him with Copernicus and Darwin, and impossible for more to agree that his psychoanalysis was "The forerunner of all of the psychotherapies." There was considerable psychology of motivation before Freud.

The book is to be praised for its many excellent presentations, discussions and warnings to the student. Its double columns, many illustrations, and emphasis upon psychology as an objective science will appeal to many. It gives a well-selected collection for the beginning student who should read some other references for somewhat different treatment of some subjects. Each chapter gives a few suggestions for further reading and a longer list of references. It contains some most recent material not found in many texts. A subject index and a name index conclude the book.

A. S. Edwards

The University of Georgia

George E. Gardner, (Ed). *Case Studies in Childhood Emotional Disabilities*, Vol. 2. New York: American Orthopsychiatric Association, 1956. Pp. 368.

The foreword of this volume (Vol. 2) indicates the continued hopes that were present in the presentation of the first volume in 1953. These were that the detailed discussions of cases "by selected Association members in the three fields of Child Psychiatry, Clinical Psychology and Psychiatric Social Work" would provide valuable teaching material for these and allied fields. The success of the first volume, requiring a second printing, has led to the publication of the present one.

The cases presented include collaborative treatment of a mother and son; planned return of child to own family; defense mechanisms of a six-year old boy; an acute separation problem; pupils psychologically absent from school; encopresis; the anal phase; brother identification in an adolescent girl; a young delinquent with psychotic mechanisms; ego treatment; a girl with atypical development; the congenitally blind; a hyperactive, destructive boy; adolescent delinquents; and a borderline psychotic child.

The degree of detail and thoroughness of study and report is at least suggested by the fact that there are only 15 reports in a 368 page book. The shortest report is 10 pages and the longest 43 pages in length, the average being a little more than 24. The amount of work in preparing these reports is more than matched by the expenditure of time in therapy (e. g., 200 visits by one boy and 125 by his mother; treatment for more than 4 years). After such intensive work and detailed reporting, the reader may come to feel that he is quite well acquainted with at least some of the cases.

Whether or not the reader is favorably inclined towards the hypotheses, theories and doctrines underlying the therapy, he may well be impressed by the skill of the therapists, the long continued and persistent efforts to give much needed help, the expensiveness of the program, and the value of the results when they are finally attained. None can study this volume without gaining valuable insight into the skills and techniques of "selected" workers in the field.

In general, the reports are simply and clearly written, and, aside from underlying theory, the student should have no trouble in understanding what he reads.

These reports do not deal with drug therapy, but considering its value and the success in teaching of the first volume, the reviewer cannot help but hope that such a book may appear, which, in similar fashion, will report case studies of the combined use of modern drug therapy and psychotherapy. Perhaps this book is the forerunner of such a volume. If drug therapy cuts time and expense, we may still be forced to agree with one of the discussants who remarked on a report of "four years of work with the mother and five years with William": "Possibly the greatest single contribution of research projects of this nature will be to prove to us that intensive treatment of very disturbed personalities is justified."

Page 368 gives the table of contents of Volume 1.

A. S. Edwards

The University of Georgia

Stanley S. Marzolf. *Psychological Diagnosis and Counseling in the Schools*. New York: Henry Holt, 1956. Pp. 401.

"This text has been written in the belief that the personnel worker in education, regardless of where and how he may wish to

specialize, should have a single, basic orientation." Emphasis is placed upon guidance. "Inherent in the conception of education as guidance is the recognition of a need for procedures for working with individuals personally, often privately, and with a high regard for those rights of individuality which characterize a free society."

Part 1 deals with the foundations of diagnosis and counseling: the sociological setting for individual methods, the evaluation of behavior and the theory of diagnosis; Part 2 takes up the case study: pertinent information, direct observation and school records, interviewing and physical measurements; Part 3 deals with testing: the characteristics of tests, administering tests, discussions of different kinds of tests, remedial teaching, counseling in theory and in action, recording and reporting.

In general, the book gives a survey of the work and duties of the personnel worker in schools, especially of the school psychologist. The problems are dealt with extensively rather than intensively, and may, in some respects, seem to overlap too much a course in testing. The discussion of testing is more extensive than that on counseling and therapy, and many readers will wish that more space had been given to the latter subjects.

A list of tests indicating "landmarks in the development of testing" is given on page 194. The Stanford-Binet revision (1916) is included but not the Terman-Merrill revision (1937). The latter is mentioned on page 241. Some of the subjects do not seem to be easily found in the index.

The sweep of subject matter seems to be rather broader than the title implies and appears to have crowded out space that might be expected to be given to actual diagnosis and counseling.

On the other hand, the wide sweep and comprehensiveness may do exactly what the author wished, namely, to orient the student in the field as a whole and the problems that are involved. Much interesting and valuable information is given. There are 17 figures and 21 tables. Many valuable suggestions are given in relation to sources of error in the work of the counselor and warnings against pitfalls and lack of proper preparation and training for the work.

The text concludes with a name index and a subject index.

A. S. Edwards

The University of Georgia

GLENN MYERS BLAIR. *Diagnostic and Remedial Teaching.* New York: Macmillan, 1956. Pp. 409.

The title of this book may raise the question, What is diagnostic teaching? The book is better than the title and deals with diagnostic testing and remedial teaching. It includes certain important subjects of special interest to teachers of elementary schools, but which are also forced upon teachers at higher levels.

The text very properly deals with how to find out what the trouble is before trying to remedy it. What are the causes of reading or of other disability, and then what are the best techniques for remedial treatment?

"Remedial teaching" in this book "refers to the two following types of activities: 1. Eliminating ineffective habits and unwholesome attitudes, and reteaching skills which have been incorrectly learned (remedying defects). 2. Teaching for the first time those habits, skills, and attitudes which have never been learned but should have been, and which are needed by pupils (developing increased competence)."

Part 1 considers diagnosis and remedial teaching of reading: the nature and results of remedial programs, determining causes of reading disability, techniques for remedial treatment, remedial programs in elementary and high schools, junior and senior, reading materials and practice exercises. Part 2 deals with arithmetic, spelling, handwriting, and fundamentals of English. Part 3 discusses improving skill in remedial teaching: making case studies and preparing for remedial teaching.

The text is enriched by 25 Tables, 20 Figures, and 16 Plates.

This is a valuable and timely book in a day when increased demand is being made for both diagnosis and remedial teaching. The importance of the book can be better understood when it is known that children are still being sent to clinics; for example, when a boy 13½ years old, is sent to the clinic with the question, "Why can't he learn to read?" an examination shows that his vision is so defective that he cannot recognize a word on a printed page. Teachers who are trained in the materials of this text should be able to detect and to diagnose many cases before the children reach the age of 13 ½ years and will learn much of remedial work which is not sufficiently well known at the present time.

Generally good advice and recommendations are given; class-

room teachers can give many tests, but the reviewer must object to the suggestion that diligent study of *Measuring Intelligence* by Terman and Merrill is a substitute for proper training in the use of any of the Binet revisions. Such examinations require thoroughly well-trained examiners, both to give the examination and to interpret results.

Many references are at the bottoms of pages, and useful bibliographies are at the ends of chapters. A well-made index (Pp. 395–409) completes the book. Much valuable material has been gathered together in this volume.

A. S. Edwards

The University of Georgia

PUBLICATIONS RECEIVED

DIDIER ANZIEU. *Le Psychodrame Analytique Chez L'Enfant.* Paris: Presses Universitaires de France, 1956, pp. 183 (paper).

ERIC BERNE. *A Layman's Guide to Psychiatry and Psychoanalysis.* New York: Simon & Schuster, 1957, pp. 320 (paper). $1.50.

FRANK H. BOWLES. *55th Report of the Director, 1956.* New York: College Entrance Examination Board, 1957, pp. 133 (paper). $.50.

JAMES M. BRADFIELD AND H. STEWARD MOREDOCK. *Measurement and Evaluation in Education.* New York: The Macmillan Co., 1957, pp. 509. $5.50.

MARGARET M. CLARK. *Left Handedness: Laterality Characteristics and Their Educational Implications.* London: University of London Press Ltd., 1957, pp. 214.

ROBERT F. DEHAAN AND ROBERT J. HAVIGHURST. *Educating Gifted Children.* Chicago: The University of Chicago Press, 1957, pp. 276. $5.00.

KATHERINE D'EVELYN. *Meeting Children's Emotional Needs: A Guide for Teachers.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1957, pp. 176. $3.75.

RUDOLF DREIKURS. *Psychology in the Classroom.* New York: Harper & Brothers, 1957, pp. 237. $3.75.

LETITIA FAIRFIELD. *Epilepsy: Grand Mal, Petit Mal, Convulsions.* New York: Philosophical Library, 1957, pp. 159. $4.75.

ARDEN N. FRANDSEN. *How Children Learn: An Educational Psychology.* New York: McGraw-Hill Book Co., 1957, pp. 546. $5.50.

CHARLES R. FOSTER. *Guidance for Today's Schools.* Boston: Ginn & Co., 1957, pp. 362. $5.00.

IRA J. GORDON. *The Teacher as a Guidance Worker.* New York: Harper & Brothers, 1956, pp. 350. $4.50.

JOSEPH KATZ, Editor. *Canadian Education Today.* Toronto: McGraw-Hill Co., Inc., 1956, pp. 243. $3.95.

PAUL LAFITTE. *The Person in Psychology: Reality or Abstraction.* New York: Philosophical Library, 1957, pp. 233. $6.00.

WESLEY P. LLOYD. *Student Personnel Services in Japan.* Washington, D.C.: American Council on Education, 1957, pp. 111. $2.00.

BERNARD J. F. LONERGAN, S.J. *Insight: A Study of Human Understanding.* New York: Philosophical Library, 1956, pp. 785. $10.00.

RUTH V. MATHENLY AND MARY TOPALIS. *Psychiatric Nursing.* Second Edition. St. Louis: C. V. Mosby Co., 1957, pp. 259. $3.50.

RAYMOND G. MCCARTHY. *Teen-Agers and Alcohol.* New Haven: Yale Center of Alcohol Studies, 1956, pp. 188. $4.00.

JOHN B. MINER. *Intelligence in the United States.* New York: Springer Publishing Co., Inc., 1957, pp. 180. $4.25.

I. P. PAVLOV. *Experimental Psychology and Other Essays.* New York: Philosophical Library, 1957, pp. 653.

JUSTIN PIKUNAS AND ALLAN P. FARRELL. *Fundamental Child Psychology.* Milwaukee: Bruce Publishing Co., 1957, pp. 259. $3.50.

DOROTHY ROGERS. *Mental Hygiene in Elementary Education.* Boston: Houghton Mifflin Co., 1957, pp. 497. $5.50.

BARUCH SPINOZA. *The Road to Inner Freedom: The Ethics.* New York: Philosophical Library, 1957, pp. 215. $3.00.

RUTH G. STRICKLAND. *The Language Arts in the Elementary School.* Second Edition. Boston: D. C. Heath and Co., 1957, pp. 464. $5.50.

DONALD E. SUPER. *The Psychology of Careers: An Introduction to Vocational Development.* New York: Harper & Bros., 1957, pp. 362. $5.75.

HARRY F. TASHMAN. *Today's Neurotic Family: A Journey into Psychoanalysis.* New York: New York University Press, 1957, pp. 214. $3.75.

HELEN M. THOMPSON. *Swierl: A Plan for Better Reading.* New York: Vintage Press, 1957, pp. 163. $3.50.

MELVIN M. TUMIN. *Segregation and Desegregation: A Digest of Recent Research.* New York: Anti-Defamation League of B'nai B'rith, 1957.

CLARK E. VINCENT. *Readings in Marriage Counseling.* New York: Thomas Y. Crowell Co., 1957, pp. 500. $4.50.

NOW AVAILABLE

Psychology and Mental Health

A report of the Institute on Education and Training for Psychological Contributions to Mental Health, held at Stanford University in August, 1955.

Edited by CHARLES R. STROTHER

Topics discussed at the Institute:

Training Needs of Psychologists in Community Mental Health

Specialization in Training

Practicum Training

Training for Therapy

Training for Research in the Mental Health Field

Problems of Departmental Organization

Price: $1.75

Order from:

American Psychological Association, Inc.
1333 Sixteenth Street, N. W.
Washington 6, D. C.

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

WM. CLARK TROW, *Chairman*
University of Michigan

STEPHEN M. COREY
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

*Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

PAUL BLOMMERS
CHESTER W. HARRIS
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street, N. W.
Washington 6, D. C.

Copyright 1958 by the American Psychological Association, Inc.

LISTENING: A CASE OF TERMINOLOGICAL CONFUSION[1]

EDNA LUE FURNESS

University of Wyoming

The effect of the auditory mode of presentation on learning and memory has been investigated in various ways since 1894. One problem which existed in the past and which still exists is that of terminology. Is the apperception of the spoken symbol a matter of hearing, of listening, or of "auding" or comprehending aurally? Another problem is that of the processes involved in the aural reception of the spoken symbol. The purpose of this paper is to show the evolution of terminology over the years and to make a differentiation of the processes involved in hearing, listening, and "auding."

At the turn of the century arose a very promising growth of monographic literature devoted solely to the psychophysics of language as such and devoted also to the conscious process involved in the apperception of spoken symbols. Among those studies is one which still remains one of the basic investigations of aural comprehension. It is a study made by the distinguished American educator, William C. Bagley.

In analyzing the apperception of the spoken sentence, Bagley (**3**) considers symbols apart from their meaning and looks at them simply as different forms of stimuli appealing to one or another of the sense departments. He shows that certain psychophysical principles condition the efficiency of perception of symbols. In the typical auditory symbol, the differences in the expressive stimuli are qualitative, temporal, and intensive. In more concrete terms, the symbolic elements differ in that they are either consonants or vowels. Consonants are distinguished by modal differences of complex nature; they are the result of different forms of adjustment of the various parts of the vocal apparatus. Vowels are distinguished by simpler qualitative differences which result from the modifications of the laryngeal clangs. These clangs are brought

[1] This article is part of a study, *A Remedial and Developmental Listening Program*, which was made possible by a grant in aid from the University of Wyoming Graduate School Council.

about by the changes in the form of the pharynx and the buccal cavity. Temporal differences are in the rapidity with which one symbolic element succeeds another symbolic element. Intensive differences are due to the fact that certain symbol elements may be emphasized and that some other symbol elements may be weakened.

Bagley continues his explanation by stating that the perception of certain letters is sufficient for the production of the verbal (auditory-kinaesthetic) idea, and for the purposes of conceptual "apperception." He answers the question concerning the character of the associative and recognitory processes that underlie the apperception of symbols by stating that, in general, the consciousness accompanying the apperception of auditory symbols is made up of sensational and affective elements. Some of these elements are aroused peripherally; others are aroused centrally in connections which vary in character with different individuals and under different conditions. "These connections," so Bagley says, "are arranged in patterns which change rapidly into one another, and are in general transitory and fleeting" (**3,** p. 125).

Bagley was concerned with the "what" of auditory language. Twenty-eight years later, Paul T. Rankin became interested in the quantification of oral expression among ordinary people. He found that the average adult spends 70% of the total waking time in some form of communication. Of that time a person spends 9% in writing activities, 16% in reading, 30% in speaking, and 45% in listening (**12,** p. 625ff.).

In the 30 years since Rankin's classic attempt to quantify the importance of auditory language, academic interest in the process of decoding spoken symbols has mounted steadily, but the serious question of what to call this form of linguistic behavior still confronts us. Rankin himself was sensitive to the inadequacies of the term "listening," and he used the word with apology and hesitation (**5,** p. 540).

Because of dissatisfaction with the vagueness of the term "listening," writers have coined a striking array of locutions which attempt to denote something more than the mere direction of attention to sound. In use are such terms as: speech reception, thoughtful hearing, getting the meaning from heard words, listening with understanding to spoken language, comprehending oral utterance, interpreting oral expression, inclining the ear cogni-

tively, hearing and understanding, listening as a language art, and aural comprehension (**8,** p. 234).

We are relatively safe from confusion with the term "aural English" so long as we are reading or writing it, but because "oral" and "aural" are homonyms in most of the major American dialects, the safer locution in an oral-aural context is the synonymous term, auditory English. Its abbreviation, "auding," has steadily gained currency in the literature of language education (**5,** p. 540).

The one who first proposed the term "auding," Don Brown, says that he has adopted it to avoid the ambiguities of such terms as "listening," "hearing," and "understanding," as well as to eliminate the awkwardness of such phrases as "getting the meaning from heard words," "listening with understanding to spoken language," or "interpreting vocal expression." As Brown uses the term, he stresses the need for more than mere awareness of perception of sound. For him, a definition of reading provides the key to a definition of "auding." For example, reading may be defined as the gross process of looking at, recognizing, and interpreting written symbols; "auding" may be defined as the gross process of listening to, recognizing, and interpreting spoken symbols (**6,** p. 128). The terms, then, which are most commonly used to indicate the apperception of the spoken language or auditory language are hearing, listening, "auding" or comprehending aurally. In the remainder of this paper the author will attempt to make a differentiation of some of the functions of these three processes.

Granted, the teaching of listening continues to be on the edge of the language arts problem, and it is still commonly assumed that hearing and listening are identical. Unfortunately, the dictionary does not offer a great deal of help in this matter of terminology. Webster says that to *listen* means "to give ear," "to hearken," "to give heed," "to hear with attention," and that to *hear* means "to have the sense or faculty of perceiving sound." Webster states further that to "hear does not necessarily imply, as *hearken* (now only poetical) and *listen* always do, attention or application; e.g., one may *hear* without *listening*, or may *listen* without *hearing*" (**11,** p. 84). Listening comprehension, then, assumes the ability to hear, but hearing does not assume the ability to listen. Hearing and listening are not synonymous. The medical doctor's concern is with hearing; the teacher's concern is with listening (**7,** p. 140).

The words "apprehend" and "comprehend" serve rather well to distinguish between hearing and listening in the sense sought in the foregoing paragraph. Remembering that apprehend means "to become aware of through the senses" and that comprehend means "to embrace or understand a thing in all its compass and extent," perhaps one could say that hearing is the apprehension of sound and that listening is the comprehension of aural symbols (**11,** p. 84).

Obviously, listening is more than hearing (**4,** p. 140). It involves following attentively the thread of a conversation, the development of an idea, the points of an argument. Like reading, it requires comprehension in terms of past experience of the listener and involves critical examination of what is heard. Whenever attention wanders, a portion of what is being presented is lost (**10,** p. 77).

Next we may consider listening and its relation to "auding." In a highly compressed 17-page chapter on "Auding" in the latest of the triennial "Language Arts" issues of the *Review of Educational Research,* Caffrey defines auding as "the process of hearing, listening to, recognizing, and interpreting or comprehending spoken language" (**9,** p. 121). The author goes on to illustrate the distinction between the first two phases, hearing and listening. He cites case studies in which patients exhibit ability to recognize the sounds of running water, crumpling paper, bells, or footsteps but fail to comprehend a word of their native tongue, or respond to their own names (**8,** p. 238). However, "many language specialists prefer to regard hearing and listening, like consciousness, as necessary CONDITIONS rather than as ASPECTS of the auding process, and restrict their concept of auding to recognizing and interpreting hearing language" (**5,** p. 541).

Poetry, *a cappella* choirs, and auto mechanics make special demands upon the ear, but, as Brown says, the relation between listening skill (control of attention to sounds) and "auding" ability (competence in decoding spoken symbols) is far from clear (**5,** p. 541). That is according to Brown. We may look to the research of other investigators. We find that of Frederick B. Agard and Harold B. Dunkel. Their research for the project called the Investigation of the Teaching of a Second Language, sponsored at the University of Chicago, 1944–1947, was used in the United States Army's language training programs. In these programs, as is well known, the

approach was predominantly oral and aural. Agard and Dunkel say: "The probability that the determining letters evoke the phonetic letter-sound images corresponding to them, and that these in turn evoke the complete verbal (auditory-kinaesthetic) images has been stated paradoxically as follows: the ability to hear more depends on the need to hear less" (**2,** p. 32). In other words, aural comprehension is dependent to a great degree upon supplementation. Agard and Dunkel's words are (**2,** p. 33): "We understand more than we hear because we fill in the gaps for ourselves without being aware of so doing. . . . It is this ability to supplement which makes the difference between auditory groping and aural comprehension."

In another work, Agard has noted that adequate "reception" or "comprehension" constitutes one-half of the act of speech communication and is composed of at least three elements (**1,** p. 292):

1. Auditory apprehension of the spoken utterance and transmission of the acoustic images to the brain for the purpose of physical or linguistic response.
2. Knowledge of the meaning to be attached to the acoustic image.
3. Awareness of certain facts or assumptions bearing on the subject of discourse but not uttered in the actual speech situations.

The evidence seems to indicate that the term "listening," as it is commonly employed today, is inadequate. The evidence seems to indicate, too, that comprehending aurally or "auding" is the more adequate term; and that "auding" consists of at least six processes: (*a*) hearing, (*b*) listening, (*c*) recognizing spoken language, (*d*) interpreting oral symbols, (*e*) supplementing meaning and knowledge of the symbols, and (*f*) being aware of facts or assumptions not uttered.

References

1. Agard, F. B. Review of Furness test of aural comprehension of Spanish. In O. K. Buros (Ed.), *Third mental measurements yearbook*. New Brunswick, N. J.: Rutgers Univer. Press, 1950.
2. Agard, F. B., & Dunkel, H. B. *An investigation of second-language teaching*. Boston: Ginn, 1948.
3. Bagley, W. C. The apperception of the spoken sentence: a study in the psychology of language. *Amer. J. Psychol.* 1900, **12,** 80–130.
4. Boyd, Gertrude, & Furness, Edna L. *Diagnostic and instructional procedures in the language arts*. Laramie, Wyoming: Curriculum and Res. Center, Univer. Wyoming, 1956.

5. Brown, D. Concepts and practices in teaching aural English. *English J.*, 1946, **35,** 540–546.
6. Brown, D. Teaching aural English. *English J.*, 1950, **39,** 128–136.
7. Brown, J. Construction of a diagnostic test of listening comprehension. *J. exp. Educ.*, 1949, **18,** 139–146.
8. Caffrey, J. G. An introduction to the auding concept. *Education*, 1949, **70,** 234–239.
9. Caffrey, J. G. Auding. *Rev. educ. Res.*, 1955, **25,** 121–138.
10. National Council of Teachers of English, Commission on the English Curriculum. *Language arts for today's children.* New York: Appleton-Century-Crofts, 1954.
11. Nichols, R. G. Listening: questions and problems. *Quart. J. Speech*, 1947, **33,** 83–86.
12. Rankin, P. T. The importance of listening ability. *English J.* 1928, **17,** 623–630.

Received March 28, 1957.

SOME RELATIONSHIPS BETWEEN SOCIOMETRIC POSITION AND LANGUAGE STRUCTURE OF YOUNG CHILDREN[1]

FRED ROSENTHAL

University of California, Berkeley

The scientific study of language has been receiving increasing attention from psychologists and educators as they have become more conscious of the pervasive nature of language in human behavior. In the present investigation, a small aspect of the general field is isolated for study. It is postulated that a relationship exists between an individual's language behavior and his position within a group. More specifically, this study is concerned with the problem of whether any significant differences can be observed in the spoken language used by children of high and low sociometric position.

A survey of the literature in this field suggests that a relationship between language and sociometric position can be demonstrated. This is derived from the following conclusions of various writers:

1. Spoken language as a social phenomenon has been related to other aspects of personality including social adjustment (**1, 2, 4, 9, 17, 18**).
2. There is some indication that individuals with poor language tend to be at the lower end of the scale in personal-social adjustment (**11, 16, 18**).
3. Social behavior as measured by sociometric methods seems to be related to other aspects of personality (**5, 8, 10, 14**).
4. There seems to be a tendency for the individuals of low sociometric status to possess the less desirable personality traits and to be more emotionally disturbed than those individuals who have higher sociometric status (**5, 13, 15, 19**).

General Procedure

Samples of spoken language from a group of second-grade children frequently chosen on a sociometric questionnaire were contrasted with the language samples of "isolated," rarely-chosen

[1] This research was carried out under the direction of Drs. Mary Cover Jones, Harold D. Carter, and Read D. Tuddenham of the University of California as partial fulfillment of the requirements for the Ph.D. degree. The reference to the dissertation appears in the bibliography for those readers who desire a fuller treatment of this problem.

children. For the language record a relatively free interview situation with an adult was employed, and each child was encouraged to talk as much as he wished.

The Interview[2]

The interviews were in two parts—in the first, six open-ended questions were asked; in the second, the child was shown six pictures and asked to tell a story about each. Each interview lasted from 15 to 30 minutes. It was recorded on a Gray Audograph machine and later transcribed to typed protocols.

Sociometric Analysis

An analysis of sociometric position was carried out on 11 second-grade classes (358 children) in six elementary schools in Oakland and Berkeley, California.[3] Each child was assessed in terms of his own class; the number of children in these classes varied from 26 to 37 (mean = 33).

In an individual interview, each child was asked to name three classmates with whom he would like to share each of four particular activities. In treating the data, all the votes a child received from his classmates were totaled so as to provide each *S* with a single score. A method developed by Bronfenbrenner (**3**) was then used to convert the raw scores to *T* scores for each class. These *T* scores can be considered comparable for the total group by evaluating them in terms of the probable number of choices which would occur by chance in a group (class) of a given size. Obtaining many more choices than would be expected by chance, given a class of a certain size, indicates high sociometric status; obtaining fewer choices than would be expected by chance indicates low sociometric status.

Children who received fewer choices in their respective classes than would be expected by chance (3% or less of the time) were placed in the "low sociometric" group; those children who received more choices than would be expected by chance (97% or more of the time) were placed in the "high sociometric" group.

---

[2] These questions and instructions are from material made available to the present writer by Dr. Walter Loban of the University of California who developed the interviews.

[3] The same schools from which the language records were obtained.

Sample

The "high" group contained 10 boys and 10 girls, and the "low" group contained 9 girls and 11 boys. In the "high" group, the ages in months ranged from 90 to 97 with the mean equal to 94 months and the standard deviation equal to 2 months; the figures for the "low" group were the same. The Kuhlman-Anderson IQ scores ranged from 92 to 112 in the "low" group with the mean and standard deviation equal to 103 and 6, respectively. In the "high" group, the IQ range was from 92 to 114 with the mean and standard deviation equal to 104 and 6, respectively. The difference between the two groups on IQ was not statistically significant. The variable of socioeconomic status (in terms of father's occupation) was controlled by having equal representation of children of various status levels in each group.

Development of Language Measures

A General Description of Language

In this investigation, language was evaluated in terms of three major characteristics: Structure, Method, and Communicative Value.

Structure. The structural elements of speech are the basic sounds (phonemes) and the connections of sounds (morphenes, words, etc.)—the building blocks out of which all languages derive their form. More specifically, this category involves a description of the materials which an individual uses in his attempt at expression. These structural elements are similar for all persons of a given language group such as English, German, French, etc.

Method. The second general category—the methodological—concerns the way in which the materials are used by a speaker in communication. Differences in style, word choice, rate of speaking, voice quality are all examples of methodological characteristics of speech.

Communicative Value. This category is closely related to the methodological. Furthermore, communicative value can be thought of as characteristic of a purposive aspect in language. In other words, as soon as verbal behavior is subjected to some nonrandom processes or methods, the language can be said to have some function—a function that is the direct expression of the rationale behind the particular methodological characteristics which are being

utilized. Thus, if the aim is communication of information, the structure must be handled according to methodological considerations which have been agreed upon by the sender and recipient. For example, when two English speaking individuals wish to exchange information, the sounds, words, sentences, etc., can not be employed in a random fashon but must follow rules which have been agreed upon or accepted (either covertly or overtly) by both if any meaning is to be transmitted.

An Operational Description of Language

In choosing the actual characteristics of language which were to be used in comparing the records of the children, two considerations were involved. Those aspects of speech were chosen which could be measured and evaluated statistically; and an attempt was made to retain meaningful, not merely measurable, features in terms of the three general characteristics discussed above. The following characteristics of language were used in the analysis:

Total Number of Words.

Communication Units. This aspect of language focuses attention on the meaningfulness, the communicative value of a message. Loban[4] in a study of children's speech has developed and used the concept of communication units:

> A word or group of words conveying a meaningful idea and having a quality of completeness about it is a communication unit. (For instance), "I know a boy with red hair." If you leave out "with red hair" you have lost the essential meaning of that sentence. "I know a boy" does not mean the same thing as "I know a boy with red hair." On the other hand, "yes" can be a unit of communication when it is an answer to "Are you going?"

Four numerical scores were derived from this category:

1. The total number communication units (CU) in a given sample of language.

2. The proportion of words used in CU to the total number of words used in a given sample of language. This proportion was defined as follows:

$$P_{CU} = \frac{\text{total of all words used in CU}}{\text{total number of words in sample}}$$

[4] This and the following quotes are from an unpublished study by Dr. Walter Loban of the University of California.

3. The mean of the number of words used in CU. This was defined as follows:

$$M_{CU} = \frac{\text{total of all words in CU}}{\text{total number of CU } (n)}$$

4. The standard deviation of the mean number of words used as CU. This was defined as follows:

$$SD_{CU} = \sqrt{\frac{\sum(X - M_{CU})^2}{n}}$$

X = the number of words in one CU

No. 1 is an indication of how many CU's are given in a sample of language; No. 2 is an indication of how much of the total language sample is valuable as communication; No. 3 indicates the average length of the CU's used in a sample of language; and No. 4 indicates how much the individual tends to vary in the lengths of the CU that he uses.

Maze. Mazes are a series of words conveying no meaning; an indication of "noise" in the message. Loban has also developed and discussed this concept giving some examples of mazes from children's speech. The mazes are in parentheses:

1. A short maze at the beginning of a communication unit and deeply related to that communication unit: "(I'm goin) ... I'm going to beat him up and kick him in the nose."

2. A short maze in the middle of the communication unit and deeply related to that communication unit: "When I was fixin' ready t'go home, mah mother called me up in the house, an' (an', an', an have to) ah have to get mah hair combed."

3. A long maze not immediately related to the communication unit, at least not deeply related: "I saw a hunter program last Sunday (an' he, an' snow time, he had to have lot hu, wah-na when he, uh, not too many dogs, he) ... and thas all I think of that picture."

In this category four numerical scores were derived in a similar manner and with similar meanings to those described in connection with communication units:

1. The total number of mazes in a given sample of language.

2. The proportion of words used in mazes to the total number of words used in a given sample of language. This was defined as $1 - P_{CU}$.

3. The mean number of words used in mazes. This was defined as follows:

$$M_m = \frac{\text{total of all words in mazes}}{\text{total number mazes } (n)}$$

4. The standard deviation of the mean number of words used as mazes. This wes defined similarly as above.

Type-Token Ratio. Johnson has discussed this measure of language extensively and described the many forms in which it can be used. He defines the TTR as:

> ...a measure of vocabulary 'flexibility' or variability, designed to indicate certain aspects of language adequacy. It expresses the ratio of different words (types) to total words (tokens) in a given language sample (**7**, p. 1).

Verb-Adjective Ratio. This measure is obtained by dividing the number of verbs (active construction) by the number of adjectives (passive construction) in a given sample of language (**2**). In this manner, a score of 100 indicates that the number of verbs and adjectives is equal; a score above 100 indicates that there are more verbs than adjectives in the particular language sample; and a score below 100 indicates that there are more adjectives than verbs. Generally, this ratio seems to be an indicator of how active and moving or how passive and inactive a sample of language tends to be (**2, 4**).

The language measures described above can be related to the three characteristics of language which have been postulated by grouping them as follows:

Structure. This category and its measures indicate the number and types of speech characteristics that these language samples contained.

1. The total number of words.
2. The total number of communication units.
3. The total number of mazes.

Method. This category gives a more qualitative description of language, and these measures indicate the variability, movement, complexity, and the richness of the sample of language.

1. The mean number of words in the communication units.
2. The mean number of words in the mazes.
3. The standard deviation of words in the communication units.
4. The standard deviation of words in the mazes.

5. The type-token ratio.
6. The verb-adjective ratio.

Communicative Value. This category consists of the proportion of the total number of words which are used as communication units (and/or $1 - P_{CU}$). This proportion indicates how much of the sample of language is meaningful communication and how much is "nonsense" from the point of view of a recipient.

Analysis

The Language Measures

In order to obtain a relatively standard situation, the analysis of language records was limited to the responses to the six pictures. The exception to this was in the computation of the total number of words which was obtained by counting all the words spoken by the child during the interview. With respect to the other aspects of language, the first 200 words and the last 200 words given in response to the pictures (or smaller segments when a child gave less than 400 words in his stories to the pictures) were analysed seperately, and the correlation of the scores obtained on the two equal segments allowed some estimate of reliability (internal consistency) of scores. A total score in each category was obtained by averaging the two scores or by adding them when interest was in a total number, e.g., the total number communication units, of mazes, of words in communication units and in mazes (in that part of the protocol used for the analysis). In each case the raw numerical scores have been transformed to normalized *T* scores for ease of statistical manipulation by the procedure outlined by Guilford (**6**).

The following ten aspects of language were drawn from the protocols of the 40 children for analysis. In the remainder of this paper they will be referred to by number:

1. The total number of words.
2. The total number of communication units (CU).
3. The percentage of the total words used in the CU.
4. The mean number of words used in the CU.
5. The standard deviation of the mean number of words used in the CU.
6. The verb-adjective ratio.
7. The mean number of words used in the mazes.

8. The total number of mazes.

9. The standard deviation of the mean number of words used in the mazes.

10. The type-token ratio.

Statistical analysis of these ten language measures on the total group of 40 children indicated the following characteristics:

1. The ten measures were converted to T scores and yielded roughly equal means (51.7 to 55.4; $\overline{X}_M = 52.4$) and standard deviations (9.53 to 11.93; $\overline{X}_{SD} = 10.63$).

2. There seemed to be some tendency for measures two to ten to be internally consistent in these records (excluding measures seven and nine) with Spearman-Brown reliabilities between two parts of the protocols ranging from .47 to .83 (see Table 1).

3. The intercorrelations of the ten measures did not indicate any significant groupings or clusters, nor any excessive degrees of overlap between these measures on the present sample (see Table 1). This indicates that the ten language categories were probably tapping different aspects of language in the present sample.

TABLE 1

PEARSON INTERCORRELATIONS AND RELIABILITIES FOR THE TEN LANGUAGE MEASURES COMPUTED FROM T SCORES

($N = 40$)

Language Measure	2	3	4	5	6	7	8	9	10	R_{11}
1	−26	−17	55**	−02	−07	−33*	−31*	21	−76**	
2		08	15	24	19	24	24	−29	41**	74
3			−06	−03	14	53**	76**	−40**	28	70
4				49**	−09	14	−23	09	−09	59
5					−22	−02	−04	00	17	47
6						11	18	−13	36*	77
7							20	−88**	−45**	15
8								09	29	83
9									−25	43

** Significance at the .01 level.

* Significance at the .05 level.

Note.—Reliability coefficients were obtained with the Spearman-Brown formula from correlations between the scores on the two segments of the protocols of nine of the language measures.

Analysis of Variance

The next step in investigating the present problem was to analyse the variance of the language measures in relation to the independent variable—sociometric status. In the present situation this was accomplished by using the three-way classification scheme presented by McNemar as case XVII (**12,** pp. 332–335); this model, with one modification, seemed to be the one that was most applicable.

The breakdown of the total degrees of freedom was in terms of R rows (individuals), B blocks (sociometric groups), and C columns (language variables). A modification was necessary because the column effect has been taken out when the ten language means had been essentially equalized by the transformation to *T* scores; and therefore, the total degrees of freedom had ten restrictions, instead of one, induced on them. Furthermore, since the individuals from block to block (high and low sociometric status) are relatively independent, i.e., even though the groups are matched for certain variables, there has been no matching by individuals—there was no R by C interaction, no B by R interaction, and no triple interaction. Table 2 summarizes the analysis of variance.

T ratios were also computed between the high and low sociometric groups for each of the ten language measures in order to allow a more intensive analysis of the language characteristics of the two groups. This is summarized in Table 3.

TABLE 2

ANALYSIS OF VARIANCE

Source	*df*	s^2	F	*p*
Individuals (Rows)	(RB − B) 38	122.8 (s_i^2)	18.1 s_b^2/s_i^2	<.001
Sociometric Groups (Blocks)	(B − 1) 1	2227.8 (s_b^2)		
Language (Columns)	0			
B × C Interaction	(B − 1) (C − 1) 9	472.8 (s_{bc}^2)	4.7 s_{bc}^2/s_h^2	<.001
Residual	B(R − 1)(C − 1) 342	100.1 (s_h^2)		
Total	(RBC − 1) 390			

TABLE 3

STANDARD ERRORS OF THE DIFFERENCES BETWEEN THE MEANS AND *t* RATIOS BETWEEN THE MEANS OF THE HIGH AND LOW SOCIOMETRIC GROUPS ON THE TEN LANGUAGE MEASURES (df = 38)

Language Measure	SDM	t	p
1	2.24	0.92	>.10
2	2.45	3.92	<.01 > .001
3	2.39	2.78	.01
4	2.21	2.42	.02
5	2.38	1.43	>.10
6	2.05	4.83	<.001
7	2.91	1.68	.10
8	2.45	1.96	<.10 > .05
9	2.22	2.21	<.05 > .02
10	2.25	2.42	.02

Results

An overall analysis of variances compared the ten measures of language simultaneously between the high and low sociometric groups. The result of this analysis indicated a nonchance difference between the high and low groups. Furthermore, it was discovered that the profiles (B × C interaction) of means for these language measures were not the same for the two extreme sociometric groups.

The significant B × C interaction in the present situation indicates that the ten language measures do not consistently differ from group to group in the same manner—that is, the high group is not superior to the low group in all measures. This interaction effect would lead to the expectations that: (*a*) not all the measures show a difference between the groups; (*b*) the differences obtained are not all of the same magnitude (equally significant); and (*c*) not all the measures differentiate these groups in the same direction, i.e., the high sociometric group having the "superior" language. In the discussion of the *t* test analysis presented below, this interpretation of the B × C interaction is upheld.

The ten language measures for the high and low sociometric groups were then compared by *t* ratios in order to obtain a more detailed picture of the differences in the two groups. This analysis indicated:

1. There was no significant difference between high and low sociometric groups in the total number of words given.

2. The high sociometric group gave a significantly smaller total number of communication units than the low sociometric group.

3. The high sociometric group gave a significantly larger percentage of words used as communication units than the low sociometric group.

4. The high sociometric group gave a significantly larger mean number of words used as communication units than the low sociometric group.

5. There was no significant difference between the high and low sociometric groups in the size of standard deviation of the mean number of words used as communication units.

6. The high sociometric group had a significantly greater verb-adjective ratio (more verbs to adjectives) than the low sociometric group.

7. The high sociometric group had a slightly lower (significant at 10%) mean number of words used as mazes than the low sociometric group.

8. The high sociometric group had a slightly smaller number of mazes than the low sociometric group.

9. The low sociometric group had a significantly larger standard deviation of the mean number of words used as mazes than the high sociometric group.

10. The high sociometric group had a significantly higher type-token ratio (more different words relative to the total number of words) than the low sociometric group.[5]

Discussion and Conclusions

Keeping in mind the hazards of generalizing from this small-scale study, the results suggest certain tendencies in regard to the language characteristics of these two groups of second-grade

[5] It was noted (in the original study and not described in this paper) that two of the variances for these ten measures (1 and 6) for high and low groups are significant. Although this must be viewed tentatively because the variances taken as a group were not significantly different (**16**), it might indicate the following possibilities: 1. The high sociometric group tends to vary more with respect to total number of words given in an interview than the low sociometric group. 2. The high sociometric individuals tend to vary more in their use of verbs relative to adjectives than the low sociometric group. This finding would tend to invalidate the t test between the means of the two groups on this measure, but since the difference between the means is highly significant (.001) there is a strong possibility that a nonchance difference remains, albeit at a lower level of significance.

children. The children of high sociometric status tend toward language characterized by a high percentage (of the total words used) of meaningful communication. Their communication is in longer units. They manifest shorter and fewer aspects of "noise" (mazes) in the language, a high usage of verbs relative to adjectives, and a greater variety of words. This group also seems to be somewhat more variable with respect to these aspects of language, especially in regard to the total number of words and the verb-adjective ratio. This latter is only suggested and could be an artifact of the tendency of this group to be higher on the language measures and, therefore, to have more opportunity for variation from zero.

The children of low sociometric status tend toward language characterized by a smaller percentage (of the total words used) of meaningful communication. Their communication is in shorter units. They use a greater number of "noise" (mazes) aspects, longer and more variable mazes, more adjectives relative to verbs, and less variety in the type of words.

There were no differences found for the two groups with respect to the total number of words used and the relative variability in length of communication units.

It is possible to state these conclusions in terms of the threefold analysis outlined earlier. Although it is not assumed that three independent aspects of language exist in terms of the measures used here, nevertheless it is meaningful to conceive of language in this way for purposes of exposition.

Structure. The available evidence does not indicate that high sociometric status is related to the tendency to use a greater quantity of language. The differences found with respect to the total number of communication units and the total number of mazes seem more related to the other aspects of language. In other words, since the groups do not differ in the total number of words used, the findings with respect to CU and mazes might be more indicative of qualitative (method and communicative value) than quantitative differences.

Method. The language of children of high sociometric status seems to be more active and moving and more varied in terms of the types of words employed. The variability in lengths of the mazes seems to be greater in the low sociometric group, but

this could be an artifact as described above. Also, the variability of the lengths of the communication units does not seem to differ for the high and low sociometric groups. Even though the individuals in the "high" group tend to give longer communication units in general, they do not vary the length of these units more than the "low" group.

Communicative value. With respect to the aspects of language which seem to be related to communicative value, the findings indicate that the children of high sociometric status are clearly superior to the children of low sociometric status. The language of "high" children is characterized by more communicative value than that of the "low" children.

In general, there was a tendency for children of high sociometric status to be differentiated from children of low sociometric status in terms of language which is more active, variable (word usage), and communicative to another individual. It seems, therefore, that children of this age and of approximately "normal" IQ tended to use an equal number of words regardless of their sociometric status with their peers in the type of interview situation used in this study. However, the more "popular" children use language of superior quality and social usefulness.

It is interesting to speculate about these findings which suggest that sociometric position is positively related to the quality and social usefulness of spoken language. Without making any assumptions about cause and effect, the following tentative "explanation" is advanced for the present results. The work of other investigators which has been mentioned seems to give some support for these views.

When an individual is interacting in a group so that he has high status (being highly chosen), he will presumably have more frequent and more diverse contact with the other members than an "isolated" individual. This in turn would make more communication channels available to the "popular" individual. Thus, the highly chosen individual will have more destinations to which he can send "messages" and a higher probability that he will receive more communication than a less popular person.

In addition, three major considerations seem to arise which favor the expectation that the verbal communication of a highly chosen, popular individual will be superior to that of an isolate.

1. The highly chosen individual has many channels open for outgoing and incoming communication. He has, therefore, many opportunities to make use of verbal communication.

2. Since learning plays an important role in the development of verbal behavior, those individuals who have greater opportunity to send and receive "messages" (more practice) will presumably have more chances to learn successful verbal habits.

3. The highly chosen individuals are also more likely to be rewarded for their attempts at communication by the interpersonal relations which result from this behavior. Also, active intercommunication between individuals will tend to encourage those habits which lead to meaningful transmission and reception—while discouraging those habits that would interfere with this task.

Summary

Twenty second-grade children of "high" sociometric position were contrasted with twenty second-grade children of "low" sociometric position on their language behavior in an interview with an adult. A conceptual schema of language was proposed in terms of structure, method, and communicative value; and the groups were contrasted on the basis of ten statistical aspects of language derived from this schema.

An attempt was made to control IQ, sex, and age; the groups were compared by an analysis of variance and *t* test techniques. The results showed significant differences between "high" and "low" groups on language behavior.

References

1. Balken, E. R., & Masserman, J. H. The language of phantasy: III. The language of the phantasies of patients with conversion hysteria, anxiety state, and obsessive-compulsive neurosis. *J. Psychol.*, 1940, **10,** 75–86.
2. Boder, D. P. The adjective-verb quotient: A contribution to the psychology of language. *Psychol. Rec.*, 1940, **3,** 310–343.
3. Bronfenbrenner, U. Measurement of sociometric status, structure, and development. *Sociom. Monogr.*, No. 6, 1945.
4. Buseman, A. *Die Sprache der Jugend als Ausdruck der Entwicklungsrhythmik.* Jena: Fischer, 1925.
5. Grossman, Beverly, & Wrighter, Joyce. The relationships between selection-rejection and intelligence, social status, and personality among sixth-grade children. *Sociometry*, 1948, **11,** 346–355.
6. Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1950.

7. Johnson, W. Studies in language behavior. I, A program of research. *Psychol. Monogr.*, 1944, **56,** No. 2 (Whole No. 255). Pp. 1–15.
8. Lindzey, G., & Borgatta, E. F. Sociometric measurement. In G. Lindzey (Ed.), *Handbook of social psychology.* Vol. I. Cambridge: Addison-Wesley, 1954.
9. Mann, Mary B. Studies in language behavior. II, The quantitative differentiation of samples of written language. *Psychol. Monogr.* 1944, **56,** No. 2 (Whole No. 255). Pp. 41–74.
10. Marks, J. B. Interests, leadership, and sociometric status among adolescents. *Sociometry,* 1954, **17,** 340–349.
11. McCarthy, Dorothea. Language development in children. In L. Carmichael, (Ed.), *Manual of child psychology.* (2nd ed.) New York: Wiley, 1954.
12. McNemar, Q. *Psychological statistics.* (2nd ed.) New York: Wiley, 1955.
13. Northway, Mary L. Outsiders: A study of the personality pattern of children least acceptable to their age mates. *Sociometry,* 1944, **3,** 10–26.
14. Northway, Mary L., Frankel, Esther B., & Potashin. Reva. Personality and sociometric status. *Sociom. Monogr.* No. 11, 1947.
15. Northway, Mary L., & Wigdor, Blossom T. Rorschach patterns related to sociometric status of school children. *Sociometry,* 1947, **10,** 186–199.
16. Rosenthal, F. Some relations between sociometric position and language structure of young children. Unpublished doctoral dissertation, Univer. of Calif., 1956.
17. Sanford, F. H. Speech and personality. *Psychol. Bull.,* 1942, **39,** 811–845.
18. Sanford, F. H. Speech and personality: A comparative case study. *Character & Pers.,* 1942, **10,** 169–198.
19. Scandrette, O. C. Classroom choice status related to scores on components of the California Test of Personality. *J. educ. Res.,* 1953, **47,** 291–296.

Received April 15, 1957.

TRANSFER OF READING TRAINING FROM NONTECHNICAL TO TECHNICAL MATERIAL

MARVIN SCHWARTZ

U. S. Naval School of Aviation Medicine, Pensacola, Florida[1]

A number of recent studies have attempted to evaluate the effectiveness of reading improvement training given college level *S*'s (e.g., **2**). These studies have been principally concerned with demonstrating that reading training results in increased reading rate, increased comprehension, and changes in academic grades. While these results are important, they do not directly answer a typical criticism leveled at reading improvement courses. The science or engineering student or the executive may express concern that, while this training may improve his ability to read a newspaper or novel, it may actually interfere with his ability to read the technical material associated with his occupation. To state it another way: should the materials employed in reading training courses be geared to the course of study of the individual student, i.e., should philosophy majors be trained on philosophical reading and mathematicians on mathematical material? The purpose of the present study was to determine if reading training utilizing relatively fictionalized, nontechnical material would transfer positively to the reading of technical material.

All beginning naval aviation students are exposed to a reading improvement course as part of their pre-flight training. It has been shown (**3**) that this training results in improved reading speed for the material utilized in training. However, since the purpose of pre-flight training is to prepare the student to be an efficient naval aviator, this reading training might better be omitted unless it can be demonstrated to aid the student in the reading of the technical material associated with his training and occupation.

Method

Subjects

The *S*'s were three classes of students in the U. S. Naval School, Pre-flight. The classes contained 27, 30, and 52 students, re-

[1] The opinions and conclusions expressed are those of the writer and not, necessarily, those of the Department of the Navy.

spectively. Because of academic difficulties in other courses resulting in failure or setback, and drops for physical reasons, 17, 17, and 40 students, respectively, completed the entire experimental procedure. Therefore, 17 from the group of 40 students were chosen, at random, to complete the experimental population. The general population represented by these three classes consists of cadets with a minimum of two years of college education. Each had been previously screened with an intelligence test, a biographical inventory, and a spatial orientation test before being accepted for flight training.

Reading Training

There are approximately 17.5 hours available for reading training during the pre-flight curriculum. Some seven of these hours are devoted to instruction in the methods and advantages of improving reading ability. The principal techniques utilized for this purpose are lectures and demonstrations with the Harvard series of reading training films. Some seven hours are devoted to actual reading practice, principally with the Reading Rate Controller. The material utilized for this practice varies from fiction and biography to difficult, but nontechnical, essay. Some examples are: *A Son of the Middle Border*, By Hamlin Garland; *The Way of a Transgressor*, by Negley Farson; and *Modern Arms and Free Men*, by Vannevar Bush. While a wide range of difficulty is represented, none of the material can be characterized as being technical in substance. Previous work (3) demonstrated that training on this type of material results in an average improvement of 88% in reading speed for this material, and, after 10 weeks of no further instruction, 90% of this improvement is retained. Comprehension, on the other hand, shows an average decrement of approximately 9% with 5% of this decrement remaining after 10 weeks.

Experimental Material

Since the material employed for the reading training is of a nontechnical nature, the object of the present experiment was to utilize as test material relatively technical selections which, while similar to material encountered in pre-flight and subsequent ground school training, would be unfamiliar to the *S*'s. Although they are given an intensive military indoctrination and training in the use of small arms, students passing through the pre-flight

curriculum are not given any instruction about naval weapons as employed in the surface fleet. Therefore, the material chosen for the present experiment consisted of three selections from a standard text on that subject, *Naval Weapons and their Uses*, published by the United States Naval Institute. These were about ships' armor, construction of naval guns, and explosives. They contained 1800, 1700, and 1725 words, respectively. The material emphasized the engineering, physics, and chemistry employed in the construction and operation of naval weapons, as well as technical nomenclature. Comprehension tests were constructed to accompany each selection. They contained a mixture of about one-half each of true-false and multiple choice items, 20 per test.

Previous work had demonstrated that reading speed could be measured with these selections with a reliability ranging from .70 to .80. The reliability of the comprehension tests was somewhat lower, .40 to .50, but of sufficient magnitude to assess group trends, as in the present experiment.

Experimental Design

Each class was tested three times: once on each of the three selections, the order of administration to each class being determined by a latin square drawn at random. The essential features of this design are such that it enables independent evaluations of the changes over sessions of the experiment, differences among the test materials, and differences among the three classes. (Differences due to order of administration of the tests were considered to be negligible and are confounded with class differences in the present experiment—cf. [**1,** p. 273] for further explication of the design.)

The initial reading and comprehension tests were given while the students were in the first week of the pre-flight program. This is an indoctrination period and was approximately three days before the reading training was begun. The posttraining test was given on the day following completion of the reading training. The retention test was administered an average of 12 weeks following the posttraining test. (The Christmas leave period caused this to vary among the classes. The minimum time between these tests was 10 weeks.) At the beginning of each test, the students were told simply that this was "a test of your ability to read and understand material of a technical nature—something you

will be expected to do throughout your respective naval careers." All tests were administered in a different setting from the one in which the reading training was conducted.

Results

Table 1 presents the mean reading speed and comprehension scores over the course of the experiment. Reading speed for technical material increased 104% between the initial and posttraining tests and showed a 92% improvement when the initial and retention tests were compared. Comprehension of the material read, however, showed a decrement. There was a 7.5% decrement comparing the initial and posttraining tests and a 5% decrement comparing the initial and retention tests. Table 2 presents the summary of the analysis of variance performed on the reading speed measure.

The principal result is the finding of a highly significant source of variance attributable to differences among experimental sessions ($F = 59.29$; 2 and 96 *df*). Applying the *t* test to the indi-

TABLE 1

MEAN READING SPEED AND COMPREHENSION SCORES

	Speed	Comprehension
Initial	193 wpm	73.0%
Posttraining	393 wpm	65.5%
Posttraining—Initial	+104%	−7.5%
Retention	372 wpm	68.0%
Retention—Initial	+ 92%	−5.0%

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE PERFORMED ON READING SPEED

Source of Variance	*df*	Variance	*F*	*p*
Between Subjects:				
Sessions × tests (Classes)	2	618.17	1.59	—
Residual	48	389.48		
Within Subjects:				
Between sessions	2	6,141.52	59.29	<.01
Between tests	2	321.60	3.12	<.05
Sessions × tests	2	176.99	1.72	—
Residual	96	103.06		

TABLE 3
SUMMARY OF ANALYSIS OF VARIANCE PERFORMED ON COMPREHENSION

Source of Variance	*df*	Variance	*F*	*p*
Between Subjects:				
Sessions × tests (classes)	2	30.88	3.32	<.05
Residual	48	9.29		
Within Subjects:				
Between sessions	2	30.77	8.45	<.01
Between tests	2	15.06	4.14	<.05
Session × tests	2	2.53	—	—
Residual	96	3.64		

vidual means indicated that the difference between the initial and posttraining sessions was significant beyond the .01 level (t = 8.93; 96 *df*) while the difference between the posttraining and the retention sessions was not significant (t = 1.05; 96 *df*). In addition, the test concerned with ships' armor was somewhat more difficult to read than the other two (t = 2.05; 96 *df*; p < .05).

Table 3 presents the summary of the analysis of variance performed on the comprehension scores. Again, the principal result is a highly significant source of variance attributable to differences among experimental sessions (F = 8.45; 2 and 96 *df*). Again, t tests showed that the difference between the initial and posttraining sessions was significant beyond the .01 level (t = 4.06; 96 *df*) while the difference between the posttraining and retention sessions was not significant (t = 1.41; 96 *df*). In addition, the selection about gun construction showed a significantly higher percentage of comprehension than did the other two tests (t = 2.49; 96 *df*; p < .05).

Discussion

The present experiment demonstrated that increased reading speed, resulting from a course of training in reading improvement, transfered to the reading of material of a different type from that employed during the training. After being trained on relatively light, nontechnical materials, the students in the present experiment demonstrated over 90% improvement in reading speed for highly technical material. For present purposes, the evaluation of the reading training in terms of the source of improvement is irrelevant. However, we can briefly state what it

probably was *not*. It appears unlikely that the improvement is completely ascribable to increased motivation, i.e., simply giving the students a pep-talk on the desirability of reading faster would not have resulted in the degree of improvement demonstrated. Further, since the training is a required part of the pre-flight curriculum for all students, the increase can not be ascribed to the fact that motivated students seeking reading instruction would benefit from *any* course of instruction. More relevantly, whatever the source of the improvement, the fact that it transfered indicates that the course had the desired effect and that the training has benefits beyond the material employed during training. Although a completely definitive test of the present problem would involve reading training on technical material, testing on technical material, and a comparison with the present program, the high reading speed transfer found in this study would suggest that it is not necessary to train readers of technical material utilizing materials from their specialties.

The finding that comprehension of the material read decreased with increased reading speed bears comment. It should be noted that this decrement is probably not a result of negative transfer from the training on nontechnical material, inasmuch as a similar decrement, of the same order of magnitude, was found for the nontechnical material after training (**3**). Thus, the decrement would appear to be due to increased speed of reading alone. It has become axiomatic in this area to say that the fast reader understands and retains more than the slow reader. The present findings would indicate that this is probably only partially true. It may be true in the sense that (*a*) the faster reader covers a greater amount of material, (*b*) less time elapses from the beginning to the end of any given selection, facillitating the integration of the material, and (*c*) the initially faster reader may also be the more intelligent reader. However, the present results indicate that speed, per se, has a deleterious rather than a beneficial effect. Nevertheless, the magnitude of the decrement demonstrated in the present instance is a small price to pay for the comparatively enormous increase in reading speed. Also, this decrement can probably be overcome by stressing that the student should not increase his speed beyond the limits of his ability to comprehend the material.

Summary

The present experiment demonstrated that reading training utilizing nontechnical reading materials for training transfered to the reading of technical material. College level *Ss* demonstrated an average of 104% improvement in reading speed for technical material after training on nontechnical material. After 10 weeks, 90% of the improvement was retained. Accompanying this increase was an average decrement of 5–7% in comprehension of the material read. It was concluded that it is not necessary to train readers of technical material utilizing materials from their specialties.

References

1. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
2. McGinnis, Dorothy. Corrective Reading: A means of increasing scholastic attainment at the college level. *J. educ. Psychol.*, 1951, **42,** 166–173.
3. Schwartz, M. An evaluation of the effectiveness of the reading training given in the US Naval School, Pre-flight. *USN Sch. Aviat. Med. Proj. Rep.*, 1957, No. NM 14 02 11 (Rep. No. 1).

Received June 18, 1957.

LEARNING ABILITY AND RETENTION AFTER LONG PERIODS[1]

D. A. Worcester

University of Wisconsin

The writer, in the summer of 1915, learned 20 100-word selections from the prose writing of Matthew Arnold and Thomas Huxley. Twelve of these were learned by the auditory method, i.e., they were read aloud to him until they could be recited, and nine were studied visually until a recitation could be made. The method allowed silent recitations during the learning process. The learner determined the rate of presentation which might vary during the course of learning. An accuracy of at least 95% was required. The time was the total elapsed time from the beginning of learning until the required degree of accuracy was achieved. Altho the goal was to learn to the point of one successful repetition, the above method might result in a higher degree of learning if there were an underestimate of the degree of memorization. One of the selections is here reproduced—one from Arnold.

> The successful future translator of Homer will have, or he cannot succeed, that true sense for his subject, and that disinterested love of it, which are, both of them, so rare in literature, and so precious; he will not be led off by any false scent; he will have an eye for the real matter, and, where he thinks he may find any indication of this, no hint will be too slight for him, no shade will be too fine, no imperfections will turn him aside—he will go before his advisor's thot and help it out with his own.

Five years later, in the summer of 1920, these selections were relearned. During this latter summer these selections and several others were used by the writer as materials to be presented to 12 other persons who memorized them. For those which he had learned previously this constituted overlearning; as to the other selections, while there was no attempt to memorize them, a high degree of familiarity was gained during the process of reading them, perhaps a total of 100 times each, to other persons.

In the summer of 1936, 21 years after the original learning,

---

[1] This paper was read before the American Psychological Association in New York, September, 1957.

the writer again studied some of these selections. He relearned 8 of the paragraphs originally learned in 1915 and relearned in 1920. He also learned 8 selections which were used by him in 1920 as materials to be learned by others—which had been, therefore, read again and again but not formally memorized. In addition he learned 4 wholly new selections, 2 by each method, from each of the authors named.

When the results of the 1936 study were reported to the American Association for the Advancement of Science, the student daily paper of the university where the writer was then employed commented upon them under the caption "Senescence No Bar To Learning." Thus encouraged, the writer undertook some more work along this line in the summer of 1956. He relearned 8 of the selections which he had learned in 1915, relearned in 1920, and relearned again in 1936. He also relearned 4 paragraphs which had been learned in 1915, relearned in 1920, but not reviewed in 1936; and 4 which had been learned in 1915, used as materials for others in 1920, but had not been relearned at that time nor reviewed since. And he learned 4 new passages. In each instance half of the selections were done by one method and half by the other.

The 4 new selections learned in July 1956 were relearned in March 1957, a lapse of time of about 8 months.

All of the selections relearned in 1920 employed the same method as that of the original learning. But in 1936 some were relearned by one method and some by the other. Altho learning was more rapid by the visual method, it appeared to make no difference in relearning time which method had been used first. No attention was given in 1956 to the initial mode of presentation.

Each of the paragraphs learned in 1915 seemed definitely familiar to the subject when relearned in 1936, and this was again true in 1956. However, some of them were felt to be much "warmer" than others.

In 1915 and in 1920 the learning exercises were done in the daytime as a part of the regular day's work. The tasks in 1936 and in 1956 were almost entirely done at night and after a full day's work at a busy summer session of the University of Nebraska. The degree to which time of day and fatigue as well as lapse of time and age may have influenced results is not known.

The writer had done almost no memoriter work in many years,

and he was especially interested to know if at his present age he could commit material by rote. There was self-assurance that he could, and he felt an eagerness to show that "He was the man he was." Whether it was from overeagerness or from some other cause, there was a definite feeling of tenseness during the study periods of 1956 and 1957 which cannot be recalled as having been present at the earlier times. There was an unwillingness to take a chance—a hesitancy to try to repeat a selection tho it seemed to be learned—but rather an attitude of "I'd better go over it once more." The very feeling of familiarity itself seemed at times to militate against good performance in the case of relearning. The content of the reading was quickly apprehended but many slips made on details—synonyms were used, for example. It is, of course, to be remembered that the increase in age was the same as the lapse of time between learning exercises.

The results of these studies are shown in the accompanying table. In terms of learning new material there seems to be an increase in time required with age. The writer doubts if time of day or fatigue had much, if any, effect. Even the somewhat jittery feeling experienced may itself be a characteristic of age. It may perhaps be noted that the learning times of 1956 were as low or lower than were the average times of those who acted as subjects of the experiment in 1920 whose ages at that time ranged from 21 to 43 (there was one who was 58). Data for the original studies are shown in references (**1, 2**).

It is striking that, altho the visual method continued to be superior to the auditory for this subject, the increase in learning time was greater by that method. A hypothesis concerning this is that reaction time, which commonly increases with age, affects the activities of reading more than those of listening. At age 67 the visual method was markedly superior to the auditory method at age 26.

The data concerning retention show several interesting things. Material relearned by the auditory method showed substantial savings over that originally learned by that method even after long lapses of time and corresponding increases in age. Material originally learned by the visual method actually required slightly more time in 1956, in most cases, for relearning than for learning. On the other hand, one notes the feeling of familiarity experienced as the selections were reread and also the very important fact that

TABLE 1

LEARNING ABILITY AND RETENTION AFTER LONG PERIODS

I Date	Age	N[a]	Average Learning Times[b] A	% over 1915	V	New Material % over 1915	
1915	26	21 A 12 V 9	13.03		6.29		
1936	47	4 A 2 V 2	14.9	14.4	7.25	15.2	
1956	67	4 A 2 V 2	16.05	23.2	10.6	68.5	
			Average Relearning Time			% Savings over 1915	
				% Savings over 1915			
1920	31	21 A 12 V 9	7.24	44.44	3.41	45.08	
1936	47	8 A 4 V 4	5.2	60.	3.9	38.	Relearning of material relearned in 1920
1936	47	8 A 4 V 4	5.4	58.6	3.9	38.	Material read many times to others but not memorized in 1920
1956	67	8 A 4 V 4	8.9	31.7	6.75	−7.3	Relearning of material relearned in 1920 and again relearned in 1936
1956	67	4 A 2 V 2	7.22	44.6	5.38	14.5	Relearned in 1920. Not reviewed 1936
1956	67	4 A 2 V 2	7.50	42.6	6.9	−9	Read to many others in 1920 but not formally relearned
1956 total		16 A 8 V 8	8.13	37.7	6.57	−4.45	
				Savings over 1956		Savings over 1956	
1957	68	4 A 2 V 2	10.	37.2	8.46	20.3	Relearning after 8 months

[a] A = by auditory presentation, V = by visual presentation.

[b] Time in minutes.

this material was relearned in about 38% less time than was required to learn new material. It has been pointed out that the increase in learning time in 1956 was greater for the visual than for the auditory method. It would appear, then, that the seemingly large loss between 1915 and 1956, especially following visual presentation, is a function more of age than of lapse of time per se. It may be said, too, that the original learning by the visual method was accomplished at a considerably greater speed for this subject than that of others who engaged in the early experiment.

In summary, the ability to memorize by rote meaningful prose was by no means lost at age 67 so far as this writer was concerned, tho the rate of learning had decreased. The decrease was more for the visual method of presentation than for the auditory, tho the visual was still the more effective method. Retention of material learned and reviewed at earlier times was substantial for that relearned by the auditory method. The results for material relearned visually were not so clear. There was probably a good deal of retention, but the factors associated with age seem to be more operative here than in the auditory mode.

It is to be hoped that others for whom early records are available will make comparisons similar to those reported here.

References

1. Worcester, D. A. Memory by visual and by auditory presentation. *J. educ. Psychol.*, 1925, **16,** 18–27.
2. Worcester, D. A. Retention after long periods. *J. educ. Psychol.*, 1923, **14,** 113–114.

Received June 25, 1957.

THE RELATIONSHIP OF CERTAIN PERSONALITY MEASURES TO EXAMINATION PERFORMANCE UNDER STRESS[1]

NEIL A. CARRIER

Southern Illinois University

One of the growing edges of recent psychology has been the study of the relationship of anxiety and behavior. Taylor (**17**), Taylor and Spence (**19**), Montague (**10**), and Lucas (**6**) have reported studies of the performance of *Ss* classified as anxious. Other investigators have tried to increase the anxiety of *Ss* by manipulating the stressful aspects of a testing situation and then measuring performance (**3, 5, 12, 16**). In general, these studies indicate that anxiety facilitates performance on simple tasks and inhibits performance on more complex activities.

Perhaps more fruitful than the mere demonstration of a relationship between anxiety and performance are attempts to determine the extent to which certain personality characteristics mediate the anxiety-performance relationship. Deese, Lazarus, and Keenan (**1**) reported an interaction between Winne Neuroticism Scale scores and conditions of shock stress under which *Ss* learned nonsense syllables. Their findings failed to appear, however, when Lazarus, Deese, and Hamilton (**4**) repeated their procedures using a more complex learning task. High anxiety now apparently inhibited rather than facilitated performance. While several other studies have tried to relate personality to performance under stress (e.g., **11, 18**), the differences reported are generally not of large magnitude. There is a need for further attempts to relate other personality measures to behavior under stress.

Motivation for the present study developed in part from interest in the work by McKeachie and others at the University of Michigan on the effect of anxiety on the examination performance of college students (**8**). While it is generally believed that small amounts of anxiety facilitate many behaviors, an assumption in

[1] Adapted from a dissertation submitted to the Doctoral Program in Social Psychology, University of Michigan. Grateful acknowledgement is extended to Dr. W. J. McKeachie for skillful guidance during this research and for helpful reading of an early draft of this manuscript.

McKeachie's work was that most students come to a course examination situation with so much anxiety about the outcome that any additional anxiety is likely to have an adverse effect on their performance. If this is true, it is possible to experimentally manipulate the stressful aspects of an examination situation and measure the corresponding effects on performance. This was done in the present study. The data were analyzed, however, according to the scores the *S*s had obtained on four previously administered personality measures.

Method

The personality measures. The personality variables selected were four that seemed to have some promise of being related to classroom examination performance. They were the SA-S personality inventory developed by Smith and reported elsewhere (**13, 14, 15**), and the need for achievement and need for affiliation measures developed by McClelland and others (**7**). The SA dimension purports to be a measure of "permeability." Persons scoring high in permeability tend to be sensitive to stimuli, impulsive, socially oriented, and low in ego strength. The S dimension ("stability" at the high end and "anxiety" at the low end) is said to reflect free-floating anxiety, nervous tension, and desurgency. The relevance of these measures to examination performance under stress should be evident. If the highly permeable student is placed in a testing situation which carries a large potential threat, his sensitivity may produce a level of anxiety which will cause a greater performance decrement than that of the relatively impermeable student. And when stress is introduced, the low stability student may show a greater performance decrement than that displayed by the more stable student.

The n Ach variable was used in view of the suggestion (**7**, pp. 110–111, 215, 271) that in males it discriminates between success-oriented (high n Ach) and failure-fearing (moderate n Ach) persons. These differently oriented persons may differ both in the kinds of attitudes they bring to a testing situation and in their abilities to handle the stressful aspects of it. The n Aff variable was used because of the implication that high n Aff persons are made more anxious by being deprived of their affiliative ties than low n Aff persons. If the high n Aff student were removed from some of his fellow students, from his teacher, and from his usual

classroom and required to take an important examination in a strange room with a strange "unaffable" teacher, he would experience more anxiety than a comparable low n Aff student.

The expected relationships of these personality variables to course examination performance under stress were formalized in a number of hypotheses:

1. High SA students in a high stress situation will score lower than high SA students in a low stress situation; no differential response to stress will be shown by moderate or low SA students.
2. Low S students in a high stress situation will score lower than similar students in a low stress situation; no differential response to stress will be shown by moderate or high S students.
3. Moderate n Ach male students in a high stress situation will score lower than similar males in a low stress situation; no differential response to stress will be shown by low or high n Ach males. (No prediction was made for female students.)
4. High n Aff students in a high stress situation will score lower than similar students in a low stress situation; no differential response to stress will be shown by moderate or low n Aff students.

Because of the uncertainty of the validity of these relatively new personality measures, it was decided that the significance of differences between groups should be tested with a two-tailed test.

Subjects. The *S*s were students in Introductory General Psychology at the University of Michigan during the spring semester, 1955. A regression equation was devised to predict the scores of the first two course examinations. The weights for these indepedent variables in the regression equation were grade-point average, American Council on Education test total score, and scores on the first two course examinations. The weights for these independent variables were computed from data on students in the same course during the previous semester.

When the predicted third examination scores were computed, they were used to form two equated groups by systematically assigning students to a low or high stress experimental situation where they took their third examination. This "matching" procedure controlled for certain intellectual and motivational factors which influence performance. The sample consisted of 125 students in each of the experimental groups. The high stress group con-

sisted of 43 male and 82 female students. The low stress group contained 60 males and 65 females. In assigning *S*s to the two groups, no thought was given to their sex. The resulting discrepancy between the number of males in the high and low stress groups must be attributed to the perversity of chance.

Stress situations. The operations by which the experimental situations were to be made differentially stressful were selected with an eye to their relevance to the personality measures employed. In the *high stress* group: the examination was given in a strange room; students were very closely proctored by a strange instructor; they were told there would be a guessing penalty in scoring the multiple-choice exam; two very difficult "practice" questions preceded the examination proper; and the instructor maintained a cold, abrupt, and rather severe demeanor throughout the examination. In the *low stress* group: the students took their examination in their usual classroom; their usual instructor was with them; they were told there would be no guessing penalty; no formal instruction sheet or practice questions were given to them; they used a special answer sheet which provided a space after each question for any comments they wished to make (**8**); and the instructor tried to maintain a warm, interested, relaxed manner and to give the impression that he was sincerely interested in assisting students to do well in every way he legitimately could.

As each student finished his examination he answered a question which asked whether he had "clutched" at any time during the examination. This question was included to determine whether differences in felt anxiety would accompany the stress conditions.

Results

A chi-square test showed significantly ($p < .02$) more clutching was reported by the high stress students than by low stress students. This provided some support for the assumed differential stressfulness of the two experimental situations.

SA: The SA dimension was broken into three groups representing high, mid, and low permeability. Test performance for each of these groups of students was then analyzed under the two experimental conditions. High SA (permeable) males and females both were significantly affected by stress (Tables 1 and 2), whereas mid and low SA students were not. These findings support the

TABLE 1
EXAMINATION SCORES OF SA MALES UNDER STRESS CONDITIONS

Stress	SA	*N*	Mean Predicted *z* Score	Mean Exam *z* Score	Relative "Gain"
Low	Lo	17	.29	.57	.28
	Mid	17	−.03	−.21	−.18
	Hi	20	−.20	.19	.39*
High	Lo	11	.10	.44	.34
	Mid	20	.04	.12	.08
	Hi	12	.13	.00	−.13*

* Significance of difference between starred values: $t = 2.26$, $p < .05$ (2 tails).

TABLE 2
EXAMINATION SCORES OF SA FEMALES UNDER STRESS CONDITIONS

Stress	SA	*N*	Mean Predicted *z* Score	Mean Exam *z* Score	Relative "Gain"
Low	Lo	22	.26	.21	−.05
	Mid	21	.12	.05	−.07
	Hi	18	−.17	.14	.31*
High	Lo	21	.28	.26	−.02
	Mid	27	.08	−.16	−.24
	Hi	33	−.13	−.38	−.25*

* Significance of difference between starred values: $t = 2.43$; $p < .02$ (2 tails).

hypothesis that stress is more detrimental to the examination performance of highly permeable students than it is to that of moderately or relatively impermeable students.

S: The hypothesis with regard to the stability dimension received only tenuous support from the data. Low stability males displayed a performance difference under stress in the predicted direction (Table 3). The t of 1.97, however, is not very reliable ($.10 > p > .05$; 2 tails). No difference approaching an acceptable level of significance was found in the case of females.

n Ach: The prediction with regard to need for achievement was unsupported. No differences in performance for high, mid, or low n Ach males under the two stress conditions were observed. Analysis of the females' data, however, showed a significant differential

TABLE 3
EXAMINATION SCORES OF S MALES UNDER STRESS CONDITIONS

Stress	S	*N*	Mean Predicted *z* Score	Mean Exam *z* Score	Relative "Gain"
Low	Lo	19	−.03	.37	.40*
	Mid	17	−.01	.03	.04
	Hi	19	.08	.15	.07
High	Lo	15	.13	.10	−.03*
	Mid	15	.03	.37	.34
	Hi	13	.08	.01	−.07

* Significance of difference between starred values: $t = 1.97$; $p < .10$ (2 tails).

TABLE 4
EXAMINATION SCORES OF n ACH FEMALES UNDER STRESS CONDITIONS

Stress	n Ach	*N*	Mean Predicted *z* Score	Mean Exam *z* Score	Relative "Gain"
Low	Lo	21	.07	.34	.27*
	Mid	14	−.01	.12	.13
	Hi	20	.01	.15	.14
High	Lo	21	−.04	−.54	−.50*
	Mid	24	−.10	−.21	−.11
	Hi	21	.17	.11	−.06

* Significance of difference between starred values: $t = 2.96$; $p < .01$ (2 tails).

reaction to stress by the low n Ach females (Table 4) such that their performance was detrimentally affected by stress.

n Aff: The hypothesis with regard to the n Aff dimension was not supported by the data. Neither males nor females displayed performance differences to stress related to their n Aff scores.

SA-S types: Since the SA and S measures correlated only .12 for males and .11 for females, another way of extracting meaning from them is to treat them as essentially noncorrelated dimensions defining four personality types—one at each of the four resulting quadrants (**14**). Median breaks were made on each dimension, and *S*s scoring within one point of either median were discarded. The remaining *S*s were classified as one of four SA-S types. Analysis of performance difference of these types showed no dif-

TABLE 5
EXAMINATION SCORES OF SA-S TYPE FEMALES UNDER STRESS CONDITIONS

Stress	Type SA	Type S	*N*	Mean Predicted z Score	Mean Exam z Score	Relative "Gain"
Low	hi	hi	14	−.06	.25	.31*
	lo	hi	13	.31	−.09	−.40
	lo	lo	10	.12	.29	.17
	hi	lo	9	−.25	−.30	−.05
High	hi	hi	19	−.15	−.41	−.26*
	lo	hi	9	.43	.36	−.07
	lo	lo	11	.22	.25	.03
	hi	lo	19	−.18	−.49	−.31

* Significance of difference between starred values: $t = 2.04$; $p < .05$ (2 tails).

TABLE 6
EXAMINATION SCORES OF ACH-AFF TYPE MALES UNDER STRESS CONDITIONS

Stress	Type Ach	Type Aff	*N*	Mean Predicted z Score	Mean Exam z Score	Relative "Gain"
Low	hi	hi	13	−.09	−.16	−.07
	hi	lo	9	.36	.44	.08
	lo	lo	17	−.07	.37	.44*
	lo	hi	15	.10	.11	.01
High	hi	hi	12	.08	.04	−.04
	hi	lo	9	.18	.22	.04
	lo	lo	10	.00	.39	.39**
	lo	hi	5	.24	.35	.11

* $t = 2.75$; $p < .02$ (2 tails).
** $t = 1.70$; $p < .15$ (2 tails).

ferential reaction to stress for males. However, female *Ss* (Table 5) who were high SA and high S types showed significantly depressed performance under stress.

Ach-Aff types: A similar type analysis was made for the n Ach and n Aff measures which correlated only .08 and .10 for males and females respectively. The males (Table 6) showed no differential reaction to stress, but an interesting observation was an apparent improvement of low n Ach-low n Aff males under both experimental conditions. The female *Ss* (Table 7) who were low n

TABLE 7
EXAMINATION SCORES OF ACH-AFF TYPE FEMALES UNDER STRESS CONDITIONS

Stress	Type		N	Mean Predicted z Score	Mean Exam z Score	Relative "Gain"
	Ach	Aff				
Low	hi	hi	16	−.12	−.01	.11
	hi	lo	10	.26	.41	.15
	lo	lo	12	.29	.82	.53*
	lo	hi	17	−.15	−.07	.08
High	hi	hi	14	.04	−.05	−.09
	hi	lo	14	.21	.19	−.02
	lo	lo	19	.07	−.32	−.25*
	lo	hi	19	−.10	−.40	−.30

* Significance of difference between starred values: $t = 2.60$; $p < .02$ (2 tails).

Ach-low n Aff showed significantly depressed performance under stress.

Discussion and Conclusions

The belief that individual performance differences in response to stress during examinations are related to detectable personality variables has received some support from this study. One conclusion suggested by our analysis is that there may be a sex difference in response to stress. The greater frequency of significant findings for females is notable. Mead (**9**) has proposed that the role of the American female is social and dependent behavior, not career accomplishment in competition with males. Field (**2**) and McClelland et al. (**7**) have reached a similar conclusion. If we can regard many female college students as persons who have not accepted academic achievement as one of their strong intrinsic motives, it may be possible to regard them as relatively poorly equipped to cope with stress in achievement situations which they perceive as only incidentally related to the gratification of their social motives. The implication here is that, if an activity is relatively unimportant or uncentral to one's dominant motives, one is poorly stimulated to develop those skills or defenses which keep the performance of this activity relatively unaffected by stress.

The writer hesitantly raises the point that, if this latter hypothesis is true, it may mean that it is easier to produce stress effects in contrived laboratory experiments than it is in experimentally manipulated "real life" situations in which the *Ss* are highly involved and to which they bring well-practiced defenses.

The SA (permeability) dimension has some promise as a means of differentiating individual performance response to stress. Since the conceptual meaning of this dimension includes sensitivity to influences from others, it is possible that the highly permeable student was most affected by the experimental manipulations which involved the behavior of the instructor during the examination. However, since the design included a number of stress manipulations, specification of the specific effects of any one is conjectural.

The finding that the high SA-high S female students showed the greatest performance decrement under stress of all the SA-S types is somewhat contrary to what might have been expected. The high SA-low S type of student has been described as "sensitive, imaginative, dependent, disorganized, uncontrolled, and cycloid" (**15**). The high SA-high S student has such characteristics as sensitivity, changeability, sociability, impulsivity, and stability. Both of these types are high on SA and, hence might be expected to show performance decrement under stress. But the high SA-high S person is relatively stable, while the high SA-low S person is relatively anxious. It might have been expected, then, that the high SA-low S female students would have performed less well under stress than the high SA-high S female students. The results indicate just the reverse. It may be that the high SA-low S female students were already so near maximal tension that the experimenter was unable to reduce their tension in the low stress situation and/or unable to increase it much in the high stress situation—and, hence, failed to get an adequate "spread" on stress for these people. In support of this interpretation is the fact that the predicted examination score of these students was the lowest of the types in both stress groups.

The n Ach variable has shown a sex interaction. While the n Ach hypothesis with regard to males has found no support, low n Ach females displayed a significant tendency to poorer performance under stress. It may be that the earlier proposal that female students experience an inhibitory conflict in academic

achievement situations is especially true of low n Ach females who may not have successfully resolved the conflict in favor of accepting academic achievement as one of their strong motives.

The n Aff variable has shown no independent discriminatory value in this study. The hypothesis with regard to n Aff is not supported.

The Ach-Aff type analysis revealed no relationship of these types to examination performance under stress for males. The apparent improvement, regardless of stress on this test relative to two previous tests, of the low n Ach-low n Aff males may indicate a kind of student who tends to "put off" study effort until rather late in a semester when the threat of failure induces intense motivation to study.

The low n Ach-low n Aff females did respond differentially under stress. This indicates that, although the n Aff variable has no predictive value when used alone, when combined with the n Ach measure it does help indentify a type of female adversely affected by examination stress.

It is recognized that a student's performance on an examination is determined by many factors. Intelligence, amount of knowledge of the subject matter, and familiarity with test-taking procedures are commonly accepted as determining influences. Our results indicate that personality characteristics of the student and the stressfulness of the testing situation must also be considered to determine performance.

Summary

The relationship of four personality variables to student performance on course examinations under experimentally manipulated stress was investigated. The findings were that the students most detrimentally affected by the stress conditions employed here were: highly permeable students, low stability males, females with low achievement motivation, females of high permeability combined with high stability, and females of low achievement motivation combined with low affiliation motivation. Also, stress had a greater detrimental effect on females than on males.

References

1. Deese, J., Lazarus, R., & Keenan, J. Anxiety, anxiety reduction, and stress in learning. *J. exp. Psychol.*, 1953, **46**, 55–60.

2. Field, W. F. The effects on thematic apperception of certain experimentally aroused needs. Unpublished doctoral dissertation, Univer. of Maryland, 1951.
3. Korchin, S. J., & Basowitz, A. Perceptual adequacy in a life stress. *J. Psychol.*, 1954, **38,** 495–502.
4. Lazarus, R., Deese, J., & Hamilton, R. Anxiety and stress in learning: the role of intraserial duplication. *J. exp. Psychol.*, 1953, **47,** 111–114.
5. Lazarus, R., & Eriksen, C. W. Effects of failure stress upon skilled performance. *J. exp. Psychol.*, 1952, **43,** 100–105.
6. Lucas, J. D. The interactive effects of anxiety, failure, and intraserial duplication. *Amer. J. Psychol.*, 1952, **65,** 59–66.
7. McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive.* New York: Appleton-Century-Crofts, 1953.
8. McKeachie, W. J., Pollie, D., & Speisman, J. Relieving anxiety in classroom examinations. *J. abnorm. soc. Psychol.*, 1955, **50,** 93–98.
9. Mead, M. *Male and female.* New York: William Morrow, 1950.
10. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45,** 91–96.
11. Osler, S. F. Intellectual performance as a function of two types of psychological stress. *J. exp. Psychol.*, 1954, **47,** 115–121.
12. Polly, S. Cognitive rigidity as a function of threat. *J. Pers.*, 1955, **23,** 346–355.
13. Smith, D. E. P., & Raygor, A. L. Verbal satiation and personality. *J. abnorm. soc. Psychol.*, 1956, **52,** 323–326.
14. Smith, D. E. P., & Wood, R. L. Noncognitive personality dimensions as indices of reading improvement at the college level. *Amer. Psychologist,* 1953, **8,** 438. (Abstract)
15. Smith, D. E. P., Wood, R. L., Downer, J. W., & Raygor, A. L. Reading improvement as a function of student personality and teaching method. *J. educ. Psychol.*, 1956, **47,** 47–59.
16. Smock, G. D. The influence of psychological stress on the "intolerance of ambiguity." *J. abnorm. soc. Psychol.*, 1955, **50,** 177–182.
17. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41,** 81–92.
18. Taylor, J. A., & Farber, I. E. The effect of failure and success as a function of ascendancy and submission. *Amer. Psychologist,* 1948, **3,** 361. (Abstract)
19. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44,** 61–64.

Received July 29, 1957.

PROBLEMS OF PROGENY RELATED TO PATERNAL EDUCATION

JAMES DRASGOW

University of Buffalo

During the process of studying psychological breakdowns among college students (**1**), it was noticed that they did not present the often observed and more obvious symptoms of anxiety or worry that so often precede breakdowns at the adult level. The overt expression of anxiety and worry appeared to follow rather than precede the breakdown episodes. In every instance, the student's college career was prematurely closed.

Another study (**2**) specifically investigated the premature closing of college careers; one of the findings was that the academic mortality rate was significantly less for students who lived in residence halls than it was for those who lived at home. While lasting longer and experiencing fewer breakdowns, the dorm students *expressed* more anxieties and worries on a simple "worry check list." This latter finding reinforced the earlier observation that those who broke down did not show manifest symptoms of anxiety prior to collapse. These previously unrelated results would appear to support the popular theoretical belief that people who are able to express their anxieties may be "draining off" whatever it might be in others that is supposed to stimulate a culminating collapse.

The present article is aimed at one of the more measurable aspects of the earlier researches. It had been found that not only did the residence hall students survive longer and express more worries, but they also had fathers with more education. For example, the majority of the fathers of the dormitory students had been exposed to at least one year of college, while the majority of the fathers of the students who lived at home had not succeeded in finishing high school; the difference was significant well beyond the .01 level.[1] This finding fostered the problems of crossvalidation and determination of degree of association, if

[1] No statistically significant relationships of discrimination or association were found for maternal education.

any, between paternal education and number of worries professed by progeny. The present project focuses on these issues.

Procedure

The folders of 50 students were drawn sequentially from the university's alphabetically ordered files. Extent of paternal education was obtained from Personnel Questionnaires that entering students completed; the number of worries was obtained from the questionnaire's "worry check list" which is reproduced in Figure 1. Each student's Minnesota Multiphasic Personality Inventory (MMPI) was classified as "stable" or "unstable" according to the following criteria: stable profiles had all scales within normal

	Considerable or Frequent Anxiety	Disturb Me Moderately	Little or No Worry
1. Poor health			
2. Fear of rats, mice or snakes			
3. Unfair treatment by teachers			
4. Harmony in the home			
5. Experiencing a sense of inferiority			
6. Being quite depressed when ignored			
7. Preferring recreation alone			
8. Being "moody"			
9. "Day-dreaming" while working			
10. Occupational possibilities			
11. Having my feelings easily hurt			
12. Finances at home			
13. Difficulty in making friends			
14. Never being asked to direct things			
15. Difficulty in concentrating			
16. Often depressed and in low spirits			
17. Slow to forget unpleasant experiences			
18. No enthusiasm about life's possibilities			
19. Worrying about little things			
20. Suspecting people of "underhanded" motives			
21. Family supervision and approval of behavior			

FIG. 1. THE WORRY CHECK LIST

TABLE 1
RELATIONSHIP BETWEEN FATHERS' EDUCATION AND THEIR COLLEGE CHILDREN'S EXPRESSED WORRIES WHEN THE TWO VARIABLES ARE CUT TO MAXIMIZE DISCRIMINATION

		Fathers' Education	
		College	No College
Worries	>3	12	13
	≦3	3	22

limits except *Mf* which is often elevated among college students, and $F - K < -12$ (**3**). All other profiles were regarded as unstable (**4**).

Results

A discriminant function analysis between the number of worries and the extent of paternal education resulted in an optimum maximizing cutoff point for discrimination as a function of worry at the three-worry level; the resulting 2×2 matrix is presented as Table 1. The cosine- pi tetrachoric correlation coefficient on this matrix is .66, and χ^2 is significant beyond .01. With these quantitative results, it may be qualitatively stated that there is indeed a sizeable and significant relationship between the number of worries expressed by a student and the number of years of education his father has had: the more educated the father is, the more worries the progeny profess. An analysis of the variance of worries among students resulted in a statistically insignificant *F*. This indicated that no one particular worry or set of worries listed in Figure 1 would serve to discriminate the children of parents with college from the children of parents without college.

The finding of $F > .05$ for worries among students was corroborated by the results obtained with the MMPI. The MMPIs, when classified as stable or unstable according to the previously indicated criteria, showed no statistically significant relationship to either number of worries indicated on the check list or to amount of paternal education.

Conclusion

A formulation from all the foregoing is that the child of a college parent may be more motivated for college than the child of a

noncollege parent since the former survives longer. The higher motivation may express itself symptomatically via worry or anxiety as indicated by the worry check list findings, but this condition in the child of the college parent is not so severe or so psychologically pathological as to manifest itself on a test of personality disturbance.

References

1. Drasgow, J. Psychological breakdown among college students. *J. higher Educ.*, 1957, **28,** 143–146.
2. Drasgow, J. Differences between college students who live in residence halls and those who live at home. *J. Educ. Res.*, in press.
3. Drasgow, J., & Barnette, W. L. The MMPI's $F - K$ in a motivated group. *J. consult. Psychol.*, 1957, **21,** 399–401.
4. Drasgow, J., & McKenzie, J. A method for predicting personality stability from transcripts. *J. counsel. Psychol.*, in press.

Received August 2, 1957.

A NEW ESTIMATE OF POLISH INTELLIGENCE[1]

BOLESLAW A. WYSOCKI

Merrimack College, Massachusetts

AND AYDIN CANKARDAS

Massachusetts Mental Health Center

In prewar Poland, applied psychology was in a comparatively rudimentary stage of development, and contact with contemporary American psychological literature was largely inadequate. Thus, the publication of Brigham's *A Study of American Intelligence* (1), which included a rather unfavorable report of Polish intelligence, produced no notable reaction among Polish psychologists. Part of the reason for this was that in Poland intelligence tests had not been widely introduced, and hence there were no scientific data to challenge Brigham's findings and conclusions. He presented a list of the national origin of the American recruits in World War I, ranked according to the percentage of each that excelled the average native American on the USA psychological examinations. Those ranged from English at the top with 67.3% to Polish at the bottom with 12.2%.

Brigham's book, which is well known in the Western Hemisphere, is often quoted with regard to national differences in mental ability. The present writers' opinion is that the conclusions drawn by Brigham have influenced American immigration quotas, and, consequently, people of Polish origin have been unduly limited in this respect. Up to the present time, no other statistical data concerning the level of Polish intelligence were available for American readers. This paper may throw a little more light on the subject of Polish intelligence. The extremely low intelligence level of Polish recruits as compared with the members of other national

[1] We wish to acknowledge the assistance of Tadeusz S. Offert, the organizer and former head of the Polish Military Psychodemographical Institute, without whose initiative and leadership these data would never have been obtained.

groups in Brigham's book is clearly contradictory to the experience of the present writers.

Method

Subjects

After the termination of the German-Polish war of 1939, a considerable part of the Polish forces, together with thousands of civilians, withdrew to the neighboring states of Hungary and Rumania. Subsequently, almost every member of this group was evacuated to France according to the prewar alternate plan formulated by the Polish government. When, later in the war, France experienced the same fate as her Polish ally, nearly all these Polish people, military and civilian, escaped to the British Isles. Following the Stalin-Sikorski Agreements of 1941, affecting over 1,000,000 Polish citizens previously deported to the Soviet Union, a second substantial group of Poles was brought to the United Kingdom. Again in 1944, after the Allied invasion of the continent, many Polish people who had been unwilling conscripts in the German army and labor battalions were reunited with the Polish national forces based in Great Britain. Finally, in 1945, after the German army had capitulated, over 100,000 members of Poland's famous General Ander's Corps, took up their station in England. These men were later joined by cilivians who had previously found refuge in the Near East. Thus in 1945, there were in Great Britain about 400,000 Poles. All classes and all parts of Poland were represented, and this group was generally regarded as a fairly representative sample of the Polish nation.

The first constructive approach to the field of applied psychology under Polish auspices was the Polish Military Psychodemographical Institute, organized in 1944, which had its central office in London, a branch office in Edinburgh, and several selection board centers distributed throughout the United Kingdom. Several officers, already possessing extensive educational backgrounds, completed a specialized training program in psychometric theory and practice at Edinburgh University under the direction of the late Godfrey H. Thomson and the late James Drever. To supplement this group of highly trained personnel, a number of psychological examiners were trained by the institute. In addition, a few of the former group continued their studies at the universities of Edinburgh and London. These, while engaged in their graduate studies, were also carrying on research projects dealing with the measurement of intelligence, apitudes, and personality factors as exhibited by the Polish population.

The Polish sample, tested by one of the writers (Wysocki) and his colleagues according to their military assignments, consisted of 15,267 men, including 651 officers. In addition, 666 adults (597 men and 69 women) were tested by the same writer in the following groups: 88 male students of a Polish university, 100 male graduates of Polish high schools, 232 male and 39 female high

school seniors, 105 men who had completed elementary school education, 36 Polish elementary school teachers, 36 Signal Corps soldiers, and 30 members of the Polish Women Navy Auxiliary Service (WAVES).

Test

Raven's Progressive Matrices (**5**) were administered to all *Ss*. The test has been widely used in England since 1938 and was considered to be particularly effective in this context because of its nonverbal character.[2]

Raven's test consists of 60 matrices from each of which a part is missing. The *Ss* has to choose the missing part to complete a pattern from six to eight alternatives drawn below each pattern. There are five sets of 12 patterns each. The first series requires mainly ability in accuracy of discrimination, but later series become progressively more difficult requiring capacity to grasp analogies, capacity to form comparisons and other logical relations. A person's intellectual capacity is assessed on the basis of his total score. The test correlates highly with Terman-Merrill scale ($r = .86$) and has a high saturation (.82) with the Spearman's g factor. The test may be administered with or without a time limit, and it can be used either as an individual test or as a group test.

The time limit of 20 minutes, as used by the British armed forces, British industrial institutions, and British schools, was implemented in all cases from which the following data were derived.

Results and Discussion

The results obtained by the writer's colleagues are reproduced in Table 1, and a few groups tested by the present writer (Wysocki) are specified in Table 2.[3]

As one would expect, Tables 1 and 2 show significant differences in test scores between officers and enlisted men, as well as among several selected groups of *Ss*. It is interesting to note that various groups of Polish nationals, as specified in Table 2, are ranked in about the same order as this test ranked similar English groups (**14**). If we look later at Tables 3 and 4, we also find (*a*) a similar hierarchical order of groups of *Ss* selected according to

[2] Coefficient of correlation with British Army S.P. 15 (verbal intelligence tests adapted by Wysocki for Polish population) was $r = .697 \pm .036$

[3] Groups 4, 6, 10, and 11 are also included in Table 1.

TABLE 1
RAVEN'S MATRICES (1938) MEAN SCORES OF POLISH OFFICERS AND ENLISTED MEN

Ss	*N*	Mean raw score	*SD*	SD_M	Mean IQ
Officers	651	45.20	7.43	.29	122.7
Enlisted men	14,616	32.50	10.39	.09	103.0

TABLE 2
RAVEN'S MATRICES (1938) MEAN SCORES OF SEVERAL GROUPS OF POLISH ADULTS

Ss	Mean raw score	Mean standard score	Mean IQ
University students (88 men)	48.64	1.85	130
High school graduates (100 men)	45.59	1.55	124
High school seniors (39 girls)	42.78	1.28	120
Military staff school students (87 officers)	42.93	1.30	120
High school seniors (232 boys)	41.00	1.11	117
Cadet-officers (564 men)	39.58	.97	116
Elementary school graduates (105 men)	39.13	.93	115
Elementary school teachers (36 men)	35.64	.59	109
Signal corps soldiers (36 men)	33.94	.43	106
Sailors (727 men)	30.71	.12	102
Infantry n.c. officers (598 men)	28.68	−.08	99
WAVES (30 women)	26.00	−.29	95

other than military criteria and (*b*) similarity with results obtained from the intelligence testing of English speaking *Ss* in Great Britain and in this country (**5, 7, 9, 12, 13, 14**). This shows that intelligence test scores of Polish *Ss* have distributions similar to those of English speaking populations and are quite comparable not only with corresponding English scores but also with those yielded by other intelligence tests in Great Britain and the United States for similar groups.

The reader, however, may wonder at the relatively low IQ levels of the Polish elementary school teachers as shown in Table 2. This may be explained by the fact that in prewar Poland elementary school teachers were trained in the five-year teachers' training schools (*Seminarium*) after having completed a seven-year elementary school. Thus, their education was on a high school level and not, as in this country or in Great Britain, on a college level. The standard of the Polish elementary school teachers'

TABLE 3

DISTRIBUTION OF IQs OF POLISH ENLISTED MEN ACCORDING TO THE LEVELS OF EDUCATION

School	Grade completed	*N*	Percentage of sample	Raven's (1938) mean raw score	Mean Stand. score	Mean IQ
Elementary schools	1	585	4.00	26.57	−.28	95
	3	1,389	9.50	26.92	−.25	96
	4 and 5	3,537	24.20	28.31	−.11	98
	6 and 7	4,443	30.40	29.22	−.02	100
Lower technical and commercial schools	10	2,558	17.50	35.00	.53	108
Secondary schools	11	1,885	12.90	38.94	.91	115
Academic schools	12 and above	219	1.50	42.55	1.26	120

education was much lower than that of those students who graduated from the Polish secondary schools, called *Gymnasium*, preparing their students for subsequent university studies. Table 2 also shows that the IQs of the army higher rank officers are lower than the IQs of high school (*Gymnasium*) graduates. This lends support to the present writers' contention that the selection to the higher officer ranks in the Polish army was not always carried out on an objective basis.

The test results obtained from enlisted men were subjected to percentage analysis with regard to several characteristics of the Polish *Ss* tested. Table 3 shows the distribution of IQ points with regard to educational levels. Here, again, one may observe a hierarchical distribution of IQs of *Ss* according to levels of education, from elementary school at the bottom to universities at the top. Again, an English intelligence test (Raven's matrices) ranked the Polish *Ss* in about the same way as it did English speaking *Ss* (**5, 8, 9**).

Table 4 shows the distribution of IQs with regard to occupational backgrounds of Polish adults. The differences between urban and rural intelligence in this country and elsewhere were widely discussed in the last two decades (**2, 3, 6, 10**). Here, in Table 4, the marked differences in mean scores between agricultural and industrial groups seem to confirm the influence of urban and rural environment in the Polish population also. Table 4 also shows significant variations of intelligence with regard to dif-

TABLE 1
RAVEN'S MATRICES (1938) MEAN SCORES OF POLISH OFFICERS AND ENLISTED MEN

Ss	*N*	Mean raw score	*SD*	SD_M	Mean IQ
Officers	651	45.20	7.43	.29	122.7
Enlisted men	14,616	32.50	10.39	.09	103.0

TABLE 2
RAVEN'S MATRICES (1938) MEAN SCORES OF SEVERAL GROUPS OF POLISH ADULTS

Ss	Mean raw score	Mean standard score	Mean IQ
University students (88 men)	48.64	1.85	130
High school graduates (100 men)	45.59	1.55	124
High school seniors (39 girls)	42.78	1.28	120
Military staff school students (87 officers)	42.93	1.30	120
High school seniors (232 boys)	41.00	1.11	117
Cadet-officers (564 men)	39.58	.97	116
Elementary school graduates (105 men)	39.13	.93	115
Elementary school teachers (36 men)	35.64	.59	109
Signal corps soldiers (36 men)	33.94	.43	106
Sailors (727 men)	30.71	.12	102
Infantry n.c. officers (598 men)	28.68	−.08	99
WAVES (30 women)	26.00	−.29	95

other than military criteria and (*b*) similarity with results obtained from the intelligence testing of English speaking *Ss* in Great Britain and in this country (**5, 7, 9, 12, 13, 14**). This shows that intelligence test scores of Polish *Ss* have distributions similar to those of English speaking populations and are quite comparable not only with corresponding English scores but also with those yielded by other intelligence tests in Great Britain and the United States for similar groups.

The reader, however, may wonder at the relatively low IQ levels of the Polish elementary school teachers as shown in Table 2. This may be explained by the fact that in prewar Poland elementary school teachers were trained in the five-year teachers' training schools (*Seminarium*) after having completed a seven-year elementary school. Thus, their education was on a high school level and not, as in this country or in Great Britain, on a college level. The standard of the Polish elementary school teachers'

TABLE 3

DISTRIBUTION OF IQS OF POLISH ENLISTED MEN ACCORDING TO THE LEVELS OF EDUCATION

School	Grade completed	*N*	Percentage of sample	Raven's (1938) mean raw score	Mean Stand. score	Mean IQ
Elementary schools	1	585	4.00	26.57	−.28	95
	3	1,389	9.50	26.92	−.25	96
	4 and 5	3,537	24.20	28.31	−.11	98
	6 and 7	4,443	30.40	29.22	−.02	100
Lower technical and commercial schools	10	2,558	17.50	35.00	.53	108
Secondary schools	11	1,885	12.90	38.94	.91	115
Academic schools	12 and above	219	1.50	42.55	1.26	120

education was much lower than that of those students who graduated from the Polish secondary schools, called *Gymnasium*, preparing their students for subsequent university studies. Table 2 also shows that the IQs of the army higher rank officers are lower than the IQs of high school (*Gymnasium*) graduates. This lends support to the present writers' contention that the selection to the higher officer ranks in the Polish army was not always carried out on an objective basis.

The test results obtained from enlisted men were subjected to percentage analysis with regard to several characteristics of the Polish *Ss* tested. Table 3 shows the distribution of IQ points with regard to educational levels. Here, again, one may observe a hierarchical distribution of IQs of *Ss* according to levels of education, from elementary school at the bottom to universities at the top. Again, an English intelligence test (Raven's matrices) ranked the Polish *Ss* in about the same way as it did English speaking *Ss* (**5, 8, 9**).

Table 4 shows the distribution of IQs with regard to occupational backgrounds of Polish adults. The differences between urban and rural intelligence in this country and elsewhere were widely discussed in the last two decades (**2, 3, 6, 10**). Here, in Table 4, the marked differences in mean scores between agricultural and industrial groups seem to confirm the influence of urban and rural environment in the Polish population also. Table 4 also shows significant variations of intelligence with regard to dif-

TABLE 4

DISTRIBUTION OF IQs OF POLISH ENLISTED MEN ACCORDING TO OCCUPATIONS

Occupation	*N*	Percentage of sample	Raven's (1938) mean raw score	Mean standard score	Mean IQ
Domestic services	44	.30	26.76	−.27	95
Agriculture	9,281	63.50	27.97	−.15	97
Industry	2,952	20.20	32.51	.29	105
Transportation	234	1.60	33.06	.34	106
Public services	219	1.50	34.25	.46	108
Commerce	59	.40	38.22	.75	112
Other occupations	1,798	12.30	40.03	1.07	116
Schools	29	.20	42.03	1.21	119

TABLE 5

RAVENS MATRICES (1938) MEAN SCORES OF ENGLISH (9) AND POLISH ADULTS ACCORDING TO AGE LEVELS

Ss	*N*	Chronological age in years					
		20	25	30	35	40	45
English[a]	5,857	44	44	42	40	38	35
Polish	3,247	31	31	29	27	25	24

[a] No time limit for English adults (9, p. 15).

ferent occupational groups—again, similar to the results obtained with the same test and the same time limit by British authors (**12, 13, 14**). Also, the age groups show marked differences in their responses to Raven's intelligence test. The decline in test scores in our sample, as shown in Table 5, may be observed already from the age of 30, and the men of 40 or 50 show a still greater decline of scores than do 30-year-old enlisted men. This characteristic decline of intelligence levels as measured by Raven's Progressive Matrices for different age groups may also be noted in Raven's test scores of British adults (**9, 12**).

The results obtained in this investigation show that scores on intelligence tests should be treated with caution. The test score alone is not necessarily a reliable index of intelligence level, particularly when one attempts to draw conclusions from the comparing of scores of different national groups. There are other factors such as education, culture, socioeconomic status, and emo-

TABLE 6

RAVEN'S MATRICES (1938) MEAN SCORES OF ENGLISH AND POLISH MEMBERS OF THE ARMED FORCES

Ss		English	Polish	*D*	SD_D	*t*
Officers	*N*	643	651			
	M	46.00	45.20	.80	.37	2.18
	SD	5.70	7.43			
	SD_M	.22	.29			
Enlisted men	*N*	15,000	14,616			
	M	34.00	32.50	1.50	.11	14.15
	SD	7.70	10.39			
	SD_M	.06	.09			

tional reactions which intelligence tests do not measure, but which may affect intelligence test scores to a marked degree. One of the present writers has received from the British War Office Selection Center (**11**) data on Raven's Progressive Matrices (1938) obtained from the testing of English officers and enlisted men. The test was administered to English adults at the same time our data from the Poles were being collected, and the same procedure of the test administration was applied. Thus, the English data were quite comparable to the data reproduced in the present investigation. Table 6 shows the critical ratio between the English and Polish officers ($t = 2.18$) statistically insignificant, and the critical ratio between the English and Polish enlisted men ($t = 14.15$) significant at the .01 level in favor of the former. Standard deviations of both groups indicate, however, that the groups of English *Ss* were more uniform than the groups of Polish *Ss*. In other words, the Polish samples were characterized by the wider range of intelligence levels which increased the critical ratio.

Let us now compare the English and Polish samples in Brigham's study of 411 and 383 *Ss* respectively (**1,** pp. 124, 126, 142). Mean combined score for the English was 14.87 and for the Poles 10.74, and the critical ratio 33.4 in favor of the former. In our investigation, there was no significant difference between the officers and not so large a difference between enlisted men of both national groups. In the majority of cases, the Polish soldiers included in the present study encountered psychological tests for the first time in their lives, whereas the English soldiers were

already familiar with testing while attending schools. Jones, discussing the same problem with regard to the American *Ss* writes: "... allowance for town familiarity with tests would give a difference of town and country of five points, which is in agreement with the eventual conclusion on the American results" (**6,** p. 5).

Fowler in his study (**4**) on 10-year-old pupils from 14 elementary schools of Detroit and Hamtramck, Michigan, has found no significant differences in mental test results of the Polish and white American nonethnic groups. Whereas in the Brooklyn, New York, area intelligence tests results of Polish children are consistently lower than those of American nonethnic children. These results point out the effect of socioeconomic status on scores of conventional intelligence tests. One has to recognize psychometric limitations of intelligence tests unless we really can develop "culture-free" tests—free from cultural learnings and unaffected by socioeconomic differences.

Summary and Conclusions

Available data concerning the level of Polish intelligence (**1**) based on a small and nonrepresentative group of Polish immigrants in the United States were felt by the writers to be inadequate. The aim of this paper was to produce recent and valid data on the intelligence of Poles. The Polish sample consisted of almost 16,000 adults. As all classes and all parts of Poland were included, this sample is the most representative and also the most recent one of adult Polish intelligence.

Raven's Progressive Matrices (1938) (**9**) were used, and the test results were subjected to percentage analysis according to different military assignments, educational levels, ages, and occupational backgrounds. It is quite a striking fact that the British-made test, when administered to various groups of Polish nationals, not only ranked various Polish groups in about the same order as this test ranked similar English groups (**5, 7, 8, 9, 11, 12, 13, 14**), but also assigned scores numerically quite comparable to corresponding English scores and indeed quite comparable to those yielded by other intelligence tests in Great Britain and the United States for comparable groups. The results reproduced in this study show clearly marked differences in intelligence levels (as measured by intelligence tests) of different groups caused by different factors such as education, occupation,

rural and industrial environment, and age levels similar to what we may observe in intelligence testing of English-speaking *S*s.

Further, the results obtained from the testing of members of the Polish armed forces were compared with the results obtained by the use of the same test on the members of English units of the British Army (**11, 14**). It appeared that the differences in mean scores between Polish and English officers were statistically negligible and the differences between English and Polish enlisted men, although statistically significant, were markedly smaller than the differences between Brigham's groups of English and Polish immigrants. This lends support to the writers' contention that, the more homogenous the groups of *S*s we compare, the smaller are the differences in intelligence levels we measure (**4**). Until really "culture-free" intelligence tests are developed—free of education, culture, and socioeconomic status influences—any comparisons of national groups based solely on intelligence tests scores should be treated with caution.

References

1. Brigham, C. *A study of American intelligence*. Princeton: Princeton Univer. Press, 1928.
2. Burt, C. *Intelligence and fertility*. London: The Eugenic Society, 1947.
3. Cattell, R. B. *The fight of our national intelligence*. London: P. S. King, 1937.
4. Dudlow, G. H. Some recent research on the Davis-Eells Games. *Sch. & Soc.*, 1956, **84**, 146–148.
5. Iscoe, I. Distribution of matrices scores of 71 college freshmen. Unpublished report, Univer. Texas, 1955.
6. Jones, H. E. Environmental handicap in mental test performance. *Univer. Calif. Publ. Psychol.*, 1932, **5**.
7. Levine, B., & Iscoe, I. A comparison of Raven's Progressive Matrices (1938) with a short form of the Wechsler-Bellevue. *Univer. Texas Report*, 1954.
8. Raven, J. C. *Progressive matrices*. London: Lewis, 1938.
9. Raven, J. C. *Guide to using progressive matrices* (1938). London: Lewis, 1956.
10. Terman, L. M., & Merrill, M. *Measuring intelligence*. London: Harper, 1937.
11. War Office Research Wing. Personal Communication. February 2, 1948.
12. Vernon, P. E. The variations of intelligence with occupation, age and locality. *Brit. J. Psychol.*, 1947, **1**, 52–63.
13. Vernon, P. E. Occupational norms for the 20-minute Progressive Matrices Test. *Occup. Psychol.*, 1949, **23**, 1.
14. Vernon, P. E. & Parry, J. B. *Personal selection in the British forces*. London: Univer. London Press, 1944.

PREDICTING INTERESTS FROM DIFFERENT KINDS OF CUES[1]

MARVIN TAYLOR

Queens College, New York

There is little doubt that effectiveness in group relations is related to the ability of the individual to perceive accurately the interests, intentions, and beliefs of other group members. Gage and his co-workers have reported positive correlations between such variables (**11, 12, 13**). Chowdrey and Newcomb found that leaders and nonleaders in various kinds of groups significantly differ in the accuracy with which they predict relevant opinions of group members (**5**). Indeed, Steiner states that one proposition underlying the rationale of teacher education is that "the more knowledge an individual has concerning the intentions, preferences, and beliefs of other persons, the more effectively he can participate in group activities with those other persons" (**15**).

Another area of research that has had a direct effect on our knowledge of social perception has been the concern with the characteristics of good and poor judges of others (**1, 3, 8, 10**). A third area of research endeavor concerned with social perception has been the attempt to find out more about the determinants of accuracy and inaccuracy of social perception (**4, 10, 11, 17**). Gronlund points out that "in the classroom, teachers are expected to judge the knowledge, attitudes, interests, and a host of other characteristics about their pupils. Although it is easier to obtain evidence in some of these areas than in others, a major share of teachers' responses are determined in part by the accuracy of their perception . . ." (**14**). Further evidence reveals that many teachers are highly inaccurate in their perceptions of students' abilities and aspirations (**7**).

Examination of the previous research suggests that one of the variables which has not been studied in any systematic way is the effect of different kinds of cues employed in social perception. For example, when a student attends a class for the first time, the teacher may perceive his motoric pattern, style of clothing,

[1] Research project completed as part of doctoral dissertation under the supervision of Francis J. Divesta, Syracuse University.

facial expressions, cleanliness, manners, verbal comments, etc., or some combination of these cues. It is probably safe to argue that social perception is mediated by both holistic and such differentiated patterns of perception. Assuming that the preceding is true, it is probable that the accuracy of an initial social perception of an individual is influenced by the mode and content of the cues available at that time. It is also probable that increments in the accuracy of social perception are influenced both by the appearance of additional cues and their order of availability. If the type of initial cues available, the appearance of additional cues, and their order of presentation could be controlled, it would seem possible to gather some valuable data about the social perception process.

Assuming that the variables mentioned above can be controlled, the folowing null hypotheses are testable:

1. There is no significant difference between a judge's accuracy in predicting interests of a stranger resulting from exposure to different types of introductory cues.
2. Successive increments of information about a stranger do not increase a judge's accuracy of prediction of a stranger's interests.
3. Final predictions of a stranger's interests are not different from earlier ones as a result of differences in the order of presentation of cues.

Method

Subjects. One hundred and eight teacher trainees (hereafter called judges) were randomly divided into six experimental sections. The subjects were predominately female; there were 18 in each section.

Experimental situation. Ten men known to be outside the peer group of the judges, but of approximately the same age group, were tested on the Minnesota Teacher Attitude Inventory (**6**) and the California F Scale (**2**). Of these ten, two, whose test scores were consistent but with marked differences between them (i.e., one may be called an authoritarian and the other a nonauthoritarian), were requested to perform certain activities (described below) before an 8-mm. movie camera and were tape recorded in an individual interview session with the experimenter. Neither of these men was informed of the activities they were expected

to engage in prior to their appearance in the movie studio. Hence it may be assumed that their behavior was spontaneous and unrehearsed. As part of the original testing program, each of the strangers had been requested to respond to a 30-item Activities Preference Index (short form).[2] The items on this instrument were selected from a pool of items shown to discriminate between authoritarian and nonauthoritarian individuals (**16**). As a test of reliability, each stranger was tested again two weeks later. No changes in response patterns were found.

The first of the filmed cue conditions involved such habitual motoric actions as walking into a room, shaking hands with another person, removing a suit jacket, and writing on a blackboard. The second filmed cue condition showed the facial expressions and hand movements of the stranger as he attempted to solve a novel and complex stylus maze problem. Finally, the third cue or verbal condition was a tape-recorded discussion of the stranger's feelings about family relations, his vocational aspirations, and his ideas on disciplining children. It is apparent that any other combination of cue conditions could have been selected, and the three described here were arbitrarily selected from the many descriptions appearing in the literature (**4, 10, 11, 17**).

The filmed and recorded cues for each stranger were presented to the groups of judges at two different 50-minute sessions. The judges were instructed to observe the stranger in one of three cue conditions and to predict his item-by-item responses on the short form Activities Preference Index. Then they were presented with a second set of cues and asked to predict again. Finally, the process was again repeated with a third set of cues. Consequently, each judge formed a first impression and made two modifications for each stranger. The two experimental sessions were three weeks apart. In order to minimize the possibilities of a practice effect, a counterbalanced design was utilized; three sections (A, B, C) saw Stranger 1 at the first session and stranger 2 at the second session. Three sections (D, E, F) saw Stranger 2 at the first session and Stranger 1 at the second. One section from each of the two counterbalanced groups (e.g., A & D, B & E, C & F) was paired and exposed to identical sequences of cues. Thus, three sequences

[2] The individual is required to respond to each statement of an activity in terms of preference or dislike for the activity. A third, or neutral category was also available as a response.

of cues were used: habitual-facial-verbal, facial-verbal-habitual, and verbal-habitual-facial.

Results

The data are reported as raw accuracy scores. These may be defined as the item-by-item correspondence between a predicted response and the actual response of the stranger on the Activities Preference Index. The analysis of variance technique was used to determine the degree of significance of mean differences. The data for each stranger were treated independently. It may be noted in Table 1 that the number of cases upon which the mean scores were based is 36 rather than 18. It will be recalled that experimental groups A and D, B and E, and C and F observed each stranger in identical sequences of cues, although the order of seeing the strangers was reversed. Since none of the mean differences between each pair of groups was found to be significant, it was decided, for the purposes of analysis, to combine the predictions into larger groupings of 36.

As might be expected, there is variation in each group of judges regardless of the type of cues utilized to form a first impression. Some individuals are more accurate in their predictions utilizing nonverbal cues, while others are more accurate using the verbal cues. On a group basis, the differences between cue condition

TABLE 1

MEANS AND STANDARD DEVIATIONS OF PREDICTIONS IN RAW ACCURACY UNITS

(N = 36 in each cell)

Cue Sequences	Stranger 1		Stranger 2	
	Mean	*SD*	Mean	*SD*
Habitual	11.6	4.3	15.1	5.0
Habitual-facial	13.7	3.7	15.1	5.6
Habitual-facial-verbal	16.2	1.6	16.1	3.0
Facial	12.4	4.0	13.3	5.0
Facial-verbal	16.2	4.0	15.9	4.4
Facial-verbal-habitual	15.2	2.7	16.8	3.9
Verbal	16.2	2.7	16.2	4.3
Verbal-habitual	14.6	2.7	17.4	4.0
Verbal-habitual-facial	15.2	2.7	17.9	3.1

TABLE 2

F TESTS OF SIGNIFICANCE BETWEEN CUE CONDITION MEANS AT EACH PREDICTION POINT

Predictions	Stranger 1			Stranger 2		
	F	*df*	*p*	*F*	*df*	*p*
One	8.44	2/105	.01	3.47	2/105	.03
Two	4.81	2/105	.01	2.43	2/105	—
Three	1.98	2/105	—	2.41	2/105	—

means on the first predictions for Stranger 1 are significant at less than the .01 level of confidence. The mean differences relevant to Stranger 2 are significant at the .03 level of confidence. The corresponding *F* ratios are 8.44 and 3.47 with 2 and 105 degrees of freedom (see Table 2). Differences in mean accuracy of second predictions for Stranger 1 are also found to be significant at the .01 level of confidence. On the other hand, the mean differences in accuracy of second and third predictions for Stranger 2 and third predictions for Stranger 1 are not significantly different.

The significant *F*s reported above were further analyzed by means of *t* tests; these results are reported in Table 3. The *t* tests were undertaken to provide a more rigorous analysis of the differential effects that each of the cue conditions might have had upon the accuracy of the first and second predictions. The results indicate that the judges using verbal cues for their initial predictions of the interests of both strangers are significantly more accurate than

TABLE 3

COMPARISON OF RAW ACCURACY MEANS AT FIRST AND SECOND PREDICTION POINTS

Cue Sequences	Stranger 1		Stranger 2	
	t	*p*	*t*	*p*
First Prediction				
Habitual-facial	1.17	—	2.22	.05
Habitual-verbal	7.74	.01	1.44	—
Facial-verbal	6.55	.01	3.76	.01
Second predictions				
Habitual, facial-verbal, habitual	2.03	.05		
Habitual, facial-facial, verbal	3.75	.01		
Facial, verbal-verbal, habitual	2.50	.01		

the judges utilizing the facial-hand cue condition. The corresponding *t*s are 6.66 and 3.76. Furthermore, the judges' predictions for Stranger 1 are significantly more accurate when they used verbal cues than when using habitual cues ($t = 7.74$). On the other hand, when predicting for Stranger 2, no significant differences in accuracy are found between judges using verbal or habitual cues. Table 3 further reveals that no other significant differences in accuracy are found when judges predicted a second and a third time for Stranger 2. On the other hand, when predicting for Stranger 1 a second time, the judges utilizing facial-verbal cues are significantly more accurate (at the .01 level of confidence) than the judges predicting from exposure to habitual-facial or verbal-habitual cues. The second predictions of judges using verbal-habitual cues are significantly more accurate (at the .05 level of confidence) than the predictions of judges utilizing habitual-facial cues.

By employing a 3 × 3 factorial design, it was also possible to explore the effects of increments of information on predictive accuracy from one prediction to another and the effects of the different sequences of cues on the accuracy of the third prediction. The results indicate that there are significant differences between each prediction, and there is also a significant interaction between the predictions and cue conditions. These relationships are found for both strangers. The differences in accuracy when using one set of cue sequences as opposed to one of the other cue sequences reveals contradictory results. An *F* ratio of 3.28 is significant at the .05 level of confidence when predicting for Stranger 2, whereas a nonsignificant *F* ratio is reported when predicting for Stranger 1. There is no readily apparent reason why there should be differences in cue sequences for one stranger and not the other.

The finding of a significant interaction for both strangers and an examination of the accuracy means (see Table 1) seems to suport the hypothesis that verbal cues are more instrumental in affecting the accuracy of predictions than either of the other two cue conditions.

Finally, the differences between successive predictions were further analyzed by means of *t* tests. Table 5 indicates that the differences in accuracy of prediction for Stranger 1 are significant at the .01 level of confidence between first and second and first and third predictions. The differences in accuracy between second

TABLE 4

F TESTS OF SIGNIFICANCE BETWEEN THREE PREDICTIONS AND THREE CUE SEQUENCES

Source of Variation	Stranger 1			Stranger 2		
	F	*df*	*p*	*F*	*df*	*p*
Between Predictions[a]	30.65	2/210	.001	9.63	2/210	.001
Interaction: Predictions × Cues	12.64	4/210	.001	2.93	4/210	.03
Between Cue sequences	2.02	2/105	—	3.28	2/105	.05

[a] Correlation between predictions was treated as suggested by Edwards (7).

TABLE 5

COMPARISON OF MEANS FOR EACH PREDICTION

Predictions	Stranger 1			Stranger 2		
	t[a]	*df*	*p*	*t*	*df*	*p*
One-Two	4.52	107	.01	2.73	107	.01
One-Three	6.04	107	.001	3.40	107	.01
Two-Three	2.12	107	.05	.59	107	—

[a] *t* test is based on correlated means.

and third predictions are significant at the .05 level of confidence. In predicting for Stranger 2, the differences in accuracy between predictions are not significant. The means and the *t* tests indicate that increasing increments of information do not have a uniform effect upon increments in the accuracy of predictions. It may be that a law of diminishing returns is in operation.

Summary

One hundred and eight teacher trainees predicted the interests of two strangers after receiving information, in different sequences, about each person. The results obtained show that the first null hypothesis does not hold. Observing a person's habitual-expressive movements and even his facial expressions when frustrated, although frequently depended on, are inadequate criteria for personality judgments as compared with listening to him speak about his feelings and aspirations. The study weakens but does not destroy the other two hypotheses tested.

References

1. Adams, H. F. The good judge of personality. *J. abnorm. soc. Psychol.*, 1927, **22,** 172–182.
2. Adorno, T. W., Frenkel-Brunswik, Else, Levinson, D. J., & Sanford, k. N. *The authoritarian personality.* New York: Harper, 1950.
3. Allport, G. W. *Personality: a psychological interpretation.* New York: Holt, 1937.
4. Bruner, J., & Taguri, R. The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology.* Vol. 2. Cambridge: Addison & Wesley, 1954. Pp. 634–650.
5. Chowdrey, K., & Newcomb, T. M. The relative abilities of leaders and non-leaders to estimate the opinions of their own group. *J. abnorm. soc. Psychol.*, 1952, **47,** 51–57.
6. Cook, W. W., Callis, R., & Leeds, C. H. *Minnesota teacher attitude inventory.* New York: Psychological Corp., 1951.
7. Cronbach, L. *Educational psychology.* New York: Harcourt, Brace, 1954.
8. Dymond, R. F. Personality and empathy. *J. consult. Psychol.*, 1950, **14,** 343–350.
9. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
10. Estes, S. G. Judging personality from expressive behavior. *J. abnorm. soc. Psychol.*, 1938, **33,** 217–236.
11. Gage, N. L. Judging interests from expressive behavior. *Psychol. Monogr.*, 1952, **66,** No. 18 (Whole No. 350).
12. Gage, N. L. Accuracy of social perception and effectiveness in interpersonal relationships. *J. Pers.*, 1953, **22,** 14–26.
13. Gage, N. L., & Suci, G. Social perception and teacher-pupil relationships. *J. educ. Psychol.*, 1951, **42,** 144–152.
14. Gronlund, N. E. The general ability to judge sociometric status: Elementary student teachers' sociometric perceptions of classmates and pupils. *J. educ. Psychol.*, 1956, 47, 147–158.
15. Steiner, I. D. Interpersonal behavior as influenced by accuracy of social perception. *Psychol. Rev.* 1955, **62,** 268–274.
16. Stern, G. G. Studies in personality typologies: The N, R, and S syndromes. Paper read at Midwest. Psychol. Ass., Chicago, May, 1953.
17. Taft, R. The ability to judge people. *Psychol. Bull.*, 1955, **52,** 1–23.

Received October 21, 1957.

SPELLING DISABILITY AND ASYNTAXIA IN A CASE INVOLVING INJURY TO THE LANGUAGE FORMULATION AREA OF THE BRAIN[1, 2]

JACK A. HOLMES

University of California

AND WILLIAM HYMAN

VA Hospital, Long Beach

In his 1955 presidential address to the APA Division of Educational Psychology, Guy T. Buswell said: "The ultimate objective, of course, is not merely to relate [psychological] research to currently accepted educational theory, but rather, to do the kind of research [in educational psychology] that will help to formulate an educational theory that is *valid*" (**1,** p. 178).

In order that learning theory in the field of educational psychology may itself be valid, however, it must be grounded in the neuroanatomy, physiology, and biochemistry of the organism that does the learning, i.e., the human being. To this end, Eames (**2**), has stressed the importance of association pathways in language disabilities, and Holmes (**6**) has worked out the relationships between the neurophysiology of the brain and the reading process. In the present paper, the writers intend to show that when, under certain circumstances, neuroanatomical evidence is combined with appropriate psychoeducational behavior, some areas of common ground may emerge.

This being the objective, the reader is asked to consider first

---

[1] Delivered before the California Educational Research Association's Annual Spring Conference, March 1957, at Santa Barbara, California.

[2] This work was done under a VA Research Fellowship at the Long Beach Veterans Hospital while the senior author was on sabbatical leave from the University of California, Berkeley. The writers express their gratitude to the VA for the use of their facilities and permission to publish case material. Likewise, the writers are indebted to J. M. Nielsen, Senior Neurological Consultant, not only for bringing this case to their attention, but also for his helpful suggestions.

the neuroclinical evidence, then the psychoeducational evidence, and finally the integration of the two.

Case History

R. J., a 35-year old, right-handed, white, male truck driver was hit with a bottle on the *left* side of his head July 14, 1956. At the receiving hospital he showed right-sided weakness, increased deep tendon reflexes on the right side, and a right Babinski sign. Angiograms, an electroencephalogram, roentgenograms, and a spinal fluid examination were reported as not remarkable, and the patient was transferred to the Veterans Administration Hospital in Long Beach. An EEG taken at this hospital was suggestive of a left-sided extra-cerebral mass—thought to be a hematoma. Before further action was taken, however, on August 26, 1956, the patient signed himself out of the hospital against medical advice.

The patient was readmitted to the VA Hospital on January 2, 1957. In the interval since his last admission, five months before, he had been bothered by transient occipital headaches, lacrimation, and diplopia. His speech had become impaired, and he complained of numbness on the right side of his face, tinnitus and pain in his right ear.

Hospital Course

A pneumoencephalogram then showed the left lateral ventricle of the brain to be moderately dilated. The neurological picture was consistent with unilateral cerebral atrophy on the left. Trephination was done bilaterally in the temporoparietal regions. Visual examination through the "burr holes" did not reveal the expected subdural hematoma.

Shortly after these diagnostic procedures had been completed, the patient began to show a gradual clinical improvement. However, two defects persisted. He had difficulty in arranging his spoken words in proper order, and he was unable to write spontaneously.

Neuroclinical Examination

The cerebral localization of the patient's present aphasia is indicated by the following clinical examination:

1. The patient is able to recognize both animate and inanimate objects; therefore, he does not have visual agnosia for objects, and we assume areas

17 and 18 to be intact. Figure 1 depicts the various areas in or on the cortex referred to during this discussion. Since R. J. is also able to recall, by revisualization, objects which he describes in color and detail, it appears that we are justified in also assuming area 19 to be intact (**7, 8**).

2. The patient is able to understand spoken commands; therefore, he shows no auditory-verbal agnosia, and we assume Wernicke's area (Broadmann's areas 41 & 42) to be intact (**7, 8**).

3. The patient is able to identify objects placed in his hand even though his eyes are closed; therefore, he manifests no tactile agnosia, and we assume the hand sensory area in the post-central gyrus to be intact, as is also the superior parietal.

4. The patient is able to carry out such a sequence of acts as is demanded by the command, "Light and smoke a cigarette." This would indicate that he does not have ideational or ideokinetic apraxia, and we assume that he does not have a diffuse cortical involvement or a deep lesion in the supramarginal gyrus (area 40) on the left side. Likewise, since he is able verbally to express his thoughts, he does not have cortical motor-pattern apraxia of the mouth, and we assume Broca's area (area 44) to be intact.

5. The patient's verbal responses manifest difficulties in language formulation. When asked, "What kind of work would you like to do?" he answered, "I want to drive for a truck." When inquiring as to the possibility of working again, he said, "If can work, how long and for work?" In response to a query regarding his previous employment, he said, "I was clean man [meaning janitor], and I was working truck, too." He, therefore, has obvious trouble with syntax and it *cannot* be assumed that area 37 on the major side is intact. Brodmann's area 37 is located, as Figure 1 shows, in the postero-lateral portion of the temporal lobe and, as first pointed out by Nielsen (**7**), is concerned with the formulation of language and with diction. This area, on the major side, coordinates word-symbol impulses from various portions of the brain and organizes them into the syntactical pattern of the coherent sentence. An injury to this language formulation area would result, then, in the patient's having difficulty with syntax, grammar, rhetoric, or style (**7, 8**).

7. The patient is able to add and subtract mentally. He can write the numbers and solve problems on paper; therefore, he shows no acalculia, and we assume his lesion has not affected the border of the interparietal sulcus or the gyri of Gratiolet (**7, 8**).

8. R. J., as mentioned above, experiences no difficulty in writing numbers. Likewise, he can copy, using the right hand, letters, words, sentences, and drawings: therefore, he does not have agraphia for *these,* and we can assume that Exner's writing center with its corresponding motor area for the hand are intact (**7, 8**).

It is somewhat misleading to speak of a "center" or specific "area" for any linguistic function as we have done, for actually any such function is so complicated that it takes many parts of the brain working together to execute such a verbal act, demon-

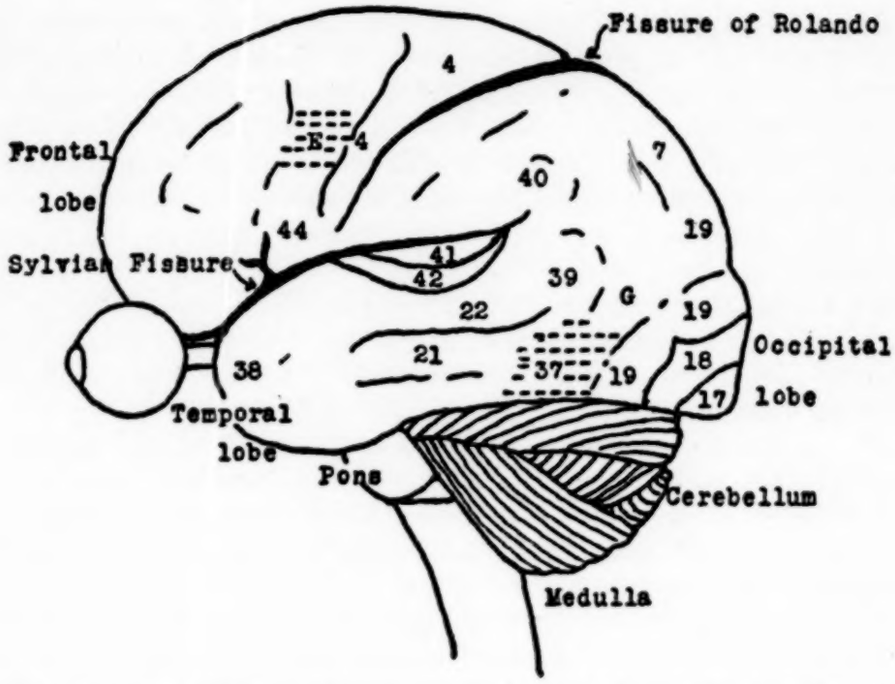

FIG. 1. BRODMANN'S AREAS INDICATED BY NUMBER (SEE TEXT FOR FUNCTIONAL INTERPRETATION).

strate a linguistic skill, or express a symbolic thought. Such a "center" as Exner's writing center, of course, could not write by itself; on the other hand, without it, the rest of the brain could not write either.

In contrast to the patient's ability to write numbers, he experiences a great deal of difficulty in trying to write the alphabet and finds it entirely impossible to write words either singly or in a sentence. As will be shown later, this is true whether the task is initiated from an oral command, a written command, or from a self-initiated desire to express himself. From these findings, it is clear that the patient *does* have agraphia for linguistic symbols. However, since the patient can write numbers and can copy words, it is assumed that the basic difficulty rests in his inability to *formulate language* (asyntaxia), and if this is correct, it would be because he has, according to Nielsen's theory (**7, 8**), a lesion in area 37 on the major side.

Psychoeducational Examination

In order to increase our understanding of the relationship of cerebral localization to psychoeducational phenomena, the following intensive differential testing was done:

1. On a Dot-Embedded Configuration Test (**3**), which assesses the ability to see a dotted outline of a symbol when it is "hidden" in a background of dots, R. J. did relatively well, scoring 78 correct out of 108 items tried.[3]

2. On a Dot-Line Configuration Test (**3**), which assesses the ability to revisualize mentally and complete a symbol when only minimal cues are given, the patient did very poorly, scoring 30 correct out of 108 items tried. This tells us the patient has trouble with "closure" when he has to supply the "closing" information from memory.

3. On a Concept Formation Test (**9**), which assesses the ability to select from 10-item-sets subgroups of four words that "belong" together because they are all associated with the same concept or idea, the patient did only fairly well, scoring 29 correct, 11 wrong, and omitting 8.

4. On a Simple Language Formulation Test (**9**), which assesses the subject's ability to select from a choice of four possible words the best one to fill a blank left in a sentence, the patient did relatively well, scoring 43 correct out of 60 items. Note that in this test recognition rather than recall is the important requirement.

5. On a Matching Sounds Test (**9**), which requires the subject visually to select paired homonyms, the patient did very well in the simple part of the test, scoring 24 out of 30; but he did very poorly in the complex section, scoring only 9 out of 30 items tried.

6. In order to test the patient's limits in his ability to formulate language, a Syntax Test (**9**) was administered in three different forms:

Form I of the Syntax Test presents a scrambled sentence as a visual stimulus and requires the subject to select, from a visual presentation of four possible arrangements of the words, the best reformulated sentence. On this form, R. J. did fairly well, scoring 19 correct out of 25.

Form II of the Syntax Test presents the same scrambled sentences but requires the subject to reformulate the word order in his own mind, remember the order, and write the best sentences as the response for each item. The patient did very poorly on this form, scoring only 7 correct out of 25 items tried. By the use of a fixed hand mirror, the patient's eye movements were observed to make many excursions from stimulus to response during the writing of each letter. In other words, not only the memory of the word order of the sentence, but also the memory for the letter order of the word did not appear to be functioning for the purpose of writing; therefore, one is forced to question this patient's ability to spell.

Form III of the Syntax Test presents the same scrambled sentence as auditory stimuli but now requires the subject to reformulate the words in his own mind, remember the word order, and say the best sentence as his response for each item. Writing is in no way involved. On this form, R. J. was not able to give a single correct response. This was not because he could not remember the words. He could, with ease, recall and repeat the sentences in their original mixed-up order, but he simply could not straighten them out!

---

[3] The standardized norms are not applicable here, because it was felt that more information could be gained by using the instruments as power rather than as speed tests. The patient was cooperative in all tests.

From the three forms of the Syntax Test, it becomes quite evident that the real trouble is in language formulation and perhaps letter formulation or *spelling from recall,* and the difficulty is not with the writing per se.

7. To this end, then, a High School Spelling Test (**3**), consisting of 35 words of increasing difficulty, was administered to the patient in four different forms:

Form I of the Spelling Test simply required the subject to recognize and encircle the one word in four that was correctly spelled. On this mutiple-choice version, the patient did surprisingly well, scoring 20 correct responses out of 35 items. Table 1 presents, in the first four columns, the choices for the five easiest items and the patient's responses to them. It should be noted that they are all correctly encircled.

Form II of the Spelling Test required the subject to write these same words from dictation. Column 5, Table 1 shows the patient's best efforts on the first five easiest items. His total score was "0."

Form III of the Spelling Test required the subject to "peck out" the letters on the typewriter in order to spell these same words from dictation. It seemed reasonable to believe that if the patient had the letters before him he might be able to select them in the proper order to fit the formulation of at least those 20 words which Form I indicated he knew how to spell correctly. However, even this proved to be an impossible task for him—his score was "0." Column 6, Table 1 gives the responses made by R. J. on the same five items.

Form IV of the Spelling Test required the subject to spell the words aloud from dictation. In spite of the fact that writing was in no way involved, R. J. was not able to spell a single word correctly. Column 7, Table 1 gives the verbal responses made by the patient to these same five words.

8. The patient did very poorly on a Phonetic Association Test (**3**), scoring 10 out of 50 items attempted.

Discussion

There is a case reported in 1953 which in many ways is strikingly parallel to the one being reported here. In other ways, however, there are marked differences.

Trowbridge (**10**) opens his article with the interesting statement: "The syndrome of loss of ability to write with preserved ability to read, is almost unknown in the field of aphasic disorders. It is the reverse of pure word blindness (inability of the patient to read what he himself has written)."

Trowbridge's patient was a 27-year old, white, navy man who had suffered a cerebral thrombosis involving branches of the left posterior cerebral artery. Six months later, the only aphasic residuals "persisting, were disorder in spelling and writing." There was no anomia, auditory-verbal, nor visual-verbal agnosia; there was no alexia, acalculia, finger agnosia, confusion of laterality, or

TABLE 1

R. J.'s RESPONSES ON FIVE EASIEST ITEMS OF THE HIGH SCHOOL SPELLING TEST UNDER FOUR DIFFERENT CONDITIONS

Recognition from visual inspection (Multiple-choice Spelling Test)				Recall from dictation		
				Writing	Typing	Oral
1	2	3	4	5	6	7[a]
1. brav	brave	brve	brev	barr	berk	da
2. dinner	dinr	diner	denr	dinrn	diinir	dinnird
3. honey	hony	huny	honny	hurn	hucy	hogny
4. pont	pount	poynt	point	rum	titc	poyn
5. theaf	theif	thief	thefe	dowd	eebc	thef

[a] Typed by examiner from patient's oral responses.

confusion of body scheme. He was able spontaneously to formulate speech conveying very complex thoughts. Unlike the present patient, Trowbridge's patient had no difficulty with either language or word formulation. However, both patients are similar in that, while they can copy words, neither can write them.

Trowbridge's patient could spell "automatically," but not when he "directed his thoughts toward spelling complex words." In this patient, the neurological examination indicated that the lesion was in the angular gyrus (area 39) on the major side. An unusual situation developed when the patient was asked to try to write with his *left* hand. While he insisted it was impossible, to everyone's surprise he did so very well! It was revealed, upon questioning, that the patient as a child *had been* left-handed. He learned to read before entering school. In school he was forcibly changed to right-handedness for writing. All his life he used the left hand for handling tools, but he had always written with his right hand. A left angular gyrus lesion then "produced an isolated loss of writing ability in the right hand, while the corresponding area on the right continued to perform normally for reading and other functions governed by this area. Writing with the left hand was possible as a result of the normal amount of skill present in this hand, together with the high level of functional development present in what would have been ordinarily the dominant angular gyrus" (**10**). R. J., on the other hand, was not able to write words with either hand.

Summary

A case report is presented in which the clinical evidence indicates that the patient had originally suffered a traumatic intracerebral hemorrhage, which had absorbed, with resulting impairment of area 37 on the major side. The relationships of area 37 to the other pertinent areas of the cortex concerned with aphasia are discussed and the possibility of their involvement eliminated or indicated. The psychoeducational significance, resulting from damage to the language formulation area (Brodmann's area 37) on the major side, is revealed by intensive differential testing.

1. Psychoeducationally, the most interesting finding is the evidence which indicates that the writing from *dictation type* of spelling ability and the *proofreader's type* of spelling ability are not one and the same ability. The significance of this finding takes on added weight when it is recalled that Holmes (**5**), in 1954, working with large samples of high school and college students was able to present statistical evidence indicating that these two types of spelling ability, while highly related, were in fact not the same. By a substrata-factor analysis, utilizing some of the same tests used on this case, Holmes was able to show that the subskills underlying the writing from dictation type of spelling ability depended more upon auditory images, while the proofreader's type of spelling ability depended more upon visual images.

2. One hesitates to draw major conclusions from a single case. However, the evidence in the present study seems to indicate that spelling ability as revealed by the ability *orally* to *spell* or to *write* words correctly from dictation calls upon a language formulation area (presumably area 37), while the proofreader's type of spelling ability to detect whether or not a word is spelled correctly by *visual* inspection does not require the use of this area. This being the case, then:

3. The most significant theoretical implication that can be drawn from this case study and the substrata-factor analysis of spelling ability is that, in the area of *spelling*, an empirical relationship has been established between the substrata factor theory (**4, 5**) and cerebral localization (**7, 8**).

4. The evidence indicates that agraphia (amnesic agraphia) can be caused by a basic inability to formulate the letters of words, even though the patient actually knows when they are spelled correctly.

5. The term *asyntaxia* is suggested as a word describing the linguistic loss resulting from injury to the language formulation area (area 37) of the brain.

References

1. Buswell, G. T. Educational theory and the psychology of learning. *J. educ. Psychol.*, 1956, **47**, 175–184.
2. Eames, T. H. Association pathways in language disabilities. *J. educ. Psychol.*, 1956, **47**, 8–10.
3. Holmes, J. A. *Diagnostic examination of visual perception and linguistic abilities.* Berkeley: Calif. Book Co., 1953.
4. Holmes, J. A. *The substrata-factor theory of reading.* Berkeley: Calif. Book Co., 1953.
5. Holmes, J. A. A substrata analysis of spelling ability for elements of auditory images. *J. exp. Educ.*, 1954, **22**, 329–350.
6. Holmes, J. A. The brain and the reading process. *Claremont Coll. Reading Conf. Yearb.*, 1957, in press.
7. Nielsen, J. M. *Agnosia, apraxia, aphasia: their value in cerebral localization.* New York: Paul B. Hoeber, 1946.
8. Nielsen, J. M. Agnosias, apraxias, speech, and aphasia. In A. B. Baker (Ed.), *Clinical neurology.* Vol. 1. New York: Paul B. Hoeber, 1955. Pp. 352–378.
9. Singer, H. *Linguistic tests: a battery for elementary schools.* Berkeley: Univer. Calif., 1956. (Mimeo.)
10. Trowbridge, W. V. Unilateral agraphia without alexia. *Bull. Los Angeles neurol. Soc.*, 1953, **18**, 153–155.

Received October 21, 1957.

Book Reviews

ROBERT W. WHITE. *Abnormal Personality.* New York: Ronald Press. 1956. Pp. 644.

How to present the rich materials of abnormal psychology in a humanized manner, intelligible to all who can read and are interested, is the challenge the author set up for himself in the first edition published in 1948 and has again adopted as his task in the second edition published in 1956. His aim, he says, is to present the subject in such a way as to emphasize the usefulness to all students of human nature, and it is because of this desire that he sees fit to include two introductory chapters, one historical and one clinical. The chief difference in the two editions is the reduction in content devoted to consideration of dreams and hypnotism and expansion of material concerned with more recent experimental studies and what their findings implicate. The author presents similar material in a more systematic manner with increased attention to developmental stages and an attempt to take an explicitly neo-Freudian position, balancing his earlier reliance on Horney with ideas drawn from Fromm and Sullivan. The case studies used include a maladjusted adolescent person, a neurotic, a delinquent, a psychotic, and a person with a progressive brain disease; and this material is used by him to advantage in a consideration of development and integration of personality. How can we find out how the difficulties of life can make one person maladjusted, another neurotic, another delinquent, another psychotic and how another passes through comparable difficulties without troubles? To bring order, system, and understanding into facts as diverse as the illustrative cases is the purpose of the whole book.

The author shows a familiarity with the literature of all topics covered and takes an eclectic attitude toward all conflicting issues. He considers the problem of therapy fairly thoroughly for an elementary text, and sees psychotherapy not as a hopelessly complex theme, but one that has to do with removing blocks in the learning process and promoting new growth where the processes are partly understood and wholly understandable. Therapies he pays a good deal of attention to are client centered, psychoanalysis, and briefer psychoanalysis. Of nondirective therapy, he says it works on a client in just one way and neglects the possibility of

using other approaches in a wise and timely fashion. He gives it as his impression that for persons with a serious neurosis, measures most patiently worked out in Freudian psychoanalysis are likely to be more promising in the long run. The basic processes considered by him as essential in all psychotherapy include (*a*) initial therapeutic relationship, (*b*) expression of feelings, (*c*) pointing out of feelings, (*d*) transference, and (*e*) behavior resulting from treatment.

There are chapters on the various mental disorders and chapters on development which are well done. The last chapter deals with the problem for society and represents a compact consideration of the meaning and implications for mankind and society of mental disease and health problems. His general conclusion is that the preventing of abnormal behavior is one in which every citizen can participate.

All in all, Dr. White has set up for himself a relatively difficult task and to a large extent has fulfilled it. To the reviewer, it is like a modernized version of Morgan's earlier efforts when less was known as given by him in *Psychology of Abnormal People*—the emphasis is on people and the knowledge given is definitely for use in understanding people. The title *Abnormal Personality* is somewhat of a misnomer since it is concerned more with people with an emphasis on diversity, rather than the thesis that there is such a thing as the abnormal personality. But that is for caption purposes only. A more honest title might be *Understanding Abnormal Personalities.*

H. MELTZER
Human Relations Research Foundation, St. Louis, Missouri

J. P. GUILFORD. *Psychometric Methods* (2nd. Ed.). New York: McGraw-Hill, 1954. Pp. 597.

When the first edition of this book appeared in 1936, specialists in educational and psychological measurement were just beginning to court mathematics and theory seriously. Thurstone headed the newly formed Psychometric Society, whose journal, *Psychometrika*, began that year. Kuder's kindred journal, *Educational and Psychological Measurement*, did not start until five years later.

Guilford's comprehensive new edition reflects changes during the eighteen-year interval and contains a great deal more material than its predecessor. New chapters on the theory of psychological

measurement and psychophysical theory build a logical basis for psychophysical and psychological scaling methods. Early in the book Guilford devotes 43 pages to explaining simple mathematics and curve fitting. He warns, however, that "it was necessary to eliminate most of the statistical treatments given in the first edition." One must go elsewhere for these, he says, preferably to his *Fundamental Statistics in Psychology and Education*.

In this second edition Guilford attempts to account for a great many phenomena of human judgment by means of Helson's concept of adaptation level. The new chapters on theories of psychological tests, reliability and validity of measures, and test development serve as vehicles for presenting recent advances. Also, much updating occurs in the chapters whose titles were altered relatively little. Finally, appendixes now include tables showing significant values of r, R, t, and F and incorporate other modifications.

Among the newcomers to Guilford's pages are Cronbach, Gulliksen, and S. S. Stevens, who have become prominent since 1936. We find little reference, however, to contemporary mathematical statisticians (none to Rao, for example). Latent structure analysis seems not to be mentioned. Coombs's recent theory and techniques get only slight notice. These and similar points are consonant with Guilford's explicit disavowal of ultra-modernity: "... this volume represents the general field of psychological measurement as of about 1950" (p. v). But one must stop somewhere!

The procedures that Guilford explains lean rather heavily upon large-sample theory, normality, and Pearson product-moment correlation, mirroring certain limitations of statistical knowledge during the early days of psychology when psychophysical and psychological scaling methods were being developed. Professional workers and graduate students in education and psychology are indebted to Guilford for the time and effort he spent revising this textbook-handbook-manual. They will find the chapters on pair(ed) comparisons, rating scales, and factor analysis particularly valuable[1].

JULIAN C. STANLEY

University of Wisconsin

[1] Prospective users should read the excellent reviews by John Keats, Clemens Johnson, and Bert Green, respectively, in *Educ. psychol. Measmt.*

ERNEST R. HILGARD. *Theories of Learning:* (2nd Ed.). New York, Appleton-Century-Crofts, 1956. Pp. 563

This edition represents a thorough revision of the widely read 1948 book. Four new chapters have been added dealing with (*a*) Freud's contributions to the psychology of learning, (*b*) the use of mathematical models, (*c*) reinforcement and drive, and (*d*) "some persisting issues." These new chapters account largely for the approximately 140 added pages, although the 1948 chapter on Wheeler's contribution has been dropped. The book carries 60 pages of bibliography, and almost half of the titles did not appear in the 1948 edition.

Hilgard's general orientation, as he developed his chapters, is much the same as it was in the earlier book. His exposition is built "around the psychologists who gave currency to the several theories"—Thorndike, Guthrie, Skinner, Hull, Tolman, Wertheimer, Koffka, Köhler, Lewin, Freud, Carr and Woodworth. Most of the chapters begin with a brief and somewhat contemporaneous history of the development of the "position," which is followed by an analysis of the major experiments, a critique of the essence of the point of view, and an estimate of its bearing upon questions such as: What are the limits of learning? What is the role of practice in learning? How important are drives and incentives, rewards and punishments? What is the place of understanding and insight? Does learning one thing help you learn something else? What happens when we remember and when we forget? Each chapter is concluded with an excellent bibliography.

Hilgard writes lucidly and his approach is generally objective, thorough, and critical. He judges Hull's work to be the best effort at "hypothetico-deductive system-making" to appear in the first half of the century (p. 121). Some cogent cautions against too much emphasis upon complicated models are suggested, and the likelihood that "a good deal of rich material from psychoanalysis has eluded the academic psychologists" (p. 307) is both lamented and explained.

Theories of Learning is written for an American audience and is probably best adapted to graduate students who are taking

(1955, **15**, 215–218), *J. Amer. stat. Assn* (1956, **51**, 413–414), and *Psychometrika* (1955, **20**, 163–165).

their first course devoted exclusively to the psychology of learning. Hilgard's personal position is eclectic and he makes a successful effort to indicate that there is something to be gained from each and all of the theoretical positions he develops. None of the systems, in his judgment, is invulnerable to criticism and the "construction of a fully satisfactory theory of learning is likely to remain for a long time an uncompleted task."

For educational psychologists who are primarily interested in finding out more about learning that can be applied in the teaching-learning situation, Hilgard's book does not have a great deal to offer. This is true even though he concludes with a discussion of "the relation between learning and the practical demands of instruction and training" (pp. 485–90). The contention is made that there are a number of important propositions about learning on which a majority of theorists would agree, and which do have clear though rather general implications for teachers. These propositions are not intended to be revolutionary. Some of them, paraphrased a bit, go like this: 1. In deciding who should learn what, the capacities of a learner are very important. 2. Motivation that is too intense may be accompanied by distracting emotional states. 3. Learning under the control of reward is usually preferable to learning under the control of punishment. 4. Learning under intrinsic motivation is preferable to learning under extrinsic motivation. 5. Tolerance for failure is best taught by providing a background of success that compensates for experienced failure. 6. There is no substitute for repeated practice in the overlearning of skills. 7. Transfer to new tasks will be better if in learning the learner can discover relationships for himself (pp. 486f).

In connection with his discussion of practical implications, Hilgard is, in this reviewer's judgment, unusually fair and moderate in his conception of the so-called conflict between pure and applied research. He neither believes that applicational research can wait until there is a pure science, nor that the existence of the pure science will automatically result in applications (p. 489). "When the relationships between pure and applied research are properly worked out, the requisite skills are of a high order. However, the labor is divided, and there is no reason why prestige should not be more equitably distributed" (p. 490). This, from so dis-

tinguished a source, gives some support to those who are disposed to believe that one of the best ways for the teacher to learn about learning is for him to experiment with his own teaching.

STEPHEN M. COREY

Teachers College, Columbia University

PAUL K. CROSSER, *The Nihilism of John Dewey*. New York: Philosophical Library, 1955.

The most interesting, as well as the most comprehensible, part of this book will be found in the first nine paragraphs of the preface. In this passage the author presents a bill of particulars with respect to the outrages which John Dewey has perpetrated on the civilization of the West.

It is made manifest that an obsession which pursued Dewey throughout his life was nothing less than the destruction of philosophy. How near he came to achieving this insane ambition can be judged by the condition in which Mr. Crosser now finds us. The American intellectual is adrift, the American artist's outlook is fragmentary, and the American teacher is "enmeshed in unrelated details." And if this were not enough, he reports also that America's perspective has been destroyed and her influence in the world diminished accordingly.

It would be a dull fellow indeed who could suppress his curiosity over how such havoc was wrought single-handed by as mild and amiable a soul as John Dewey is reputed to have been. It is likely, therefore, that many readers will turn with anticipation from the preface to the body of the text. Those who are not in the habit of reading prefaces will be in on the secret from the beginning—that this is a monument to dullness. At least the preface reader can enjoy for a few moments his illusion that at long last the anatomy of destruction is to be laid bare.

The body of the work—232 pages of it—is devoted to shredding large chunks of three of Dewey's major works: *Logic, The Theory of Inquiry, Art as Experience*, and *Experience and Education*. The technique of what its author optimistically calls "this critical analysis" is to cite a sentence or phrase from Dewey and then envelop it in the heaviest sarcasm to be found outside a sophomore bull session. Sarcasm, particularly when it is prolonged for 232 pages, is difficult for even a good writer to manage—and

in truth Mr. Crosser is not to be numbered among these. Sarcasm, used sparingly and with finesse, can be deadly to an opponent. But here it serves only to obscure the argument and to create an atmosphere of petty contentiousness, which would be distasteful even to one who was sympathetic to the argument itself.

The analysis is directed towards certain aspects of Dewey's theory of knowledge, his theory of art, and his theory of education. The author maintains that in all of these Dewey's nihilism is apparent. The chapter titles are indicative of the tenor of the exposition: "The Degradation of Knowledge," "The Dissipation of the Sense of Beauty," and "The Exclusion of Schools."

The blunt instrument which was used in the initial crime against the West was the kind of "cognitive indeterminableness" which is alleged to be the root of Dewey's villainy. Through it the natural and social sciences have been struck down. Subject and object, cause and effect disappear into a cognitive limbo. The inevitable result is *extreme relativism*. (Crosser always puts this term in italics lest we forget that it is a *bad thing*.)

The rest of catastrophe follows hard on this. Objects of art are consigned to outer darkness, followed quickly by the artist. The sense of beauty is lost and the very definition of esthetics becomes obscured in a miasma of *extreme relativism*.

With respect to the state of formal education the situation is, if anything, worse. The price we have paid for our cognitive undeterminableness is that the school as an institution is fast disappearing. Our concept of education as formal learning of organized knowledge is all but forgotten. In our *extreme relativism* we are even unable to discern what is educative experience and what is not. The decline of the West is all but accomplished.

It is too early to discern how telling this blow will be against the philosophy of nihilism, but it seems likely that Dewey's reputation has already been diminished in one respect. It has long been advocated in some quarters that Dewey had the worst prose style to be found in five generations. If this was true (and there are many who deny it), it is hardly true now. Dewey has met his match, and in support of this thesis the following passage from page 119 is offered in evidence:

"The cognitive violence which the removal of art objects as art objects as an analytical factor from the kind of art theory he is about to evolve is of such serious nature that the return of

the art object into the fold of inquiry, after it had been left out of consideration in the evolvement of the basic principles of that inquiry, cannot restore to that kind of inquiry the character of an inquiry into the arts."

G. MAX WINGO

University of Michigan

WILLIAM E. HENRY, *The Analysis of Fantasy*, New York: John Wiley & Sons, Inc., London: Chapman & Hall, Limited. 1956, Pp. 305.

The Analysis of Fantasy is a consideration of the use of thematic apperception technique in the analysis of personality. The title is smarter as is, and to some extent can be said to be justified since the thematic apperception is one of the relatively effective methods of getting at fantasy for the sake of interpreting symbolic behavior. The material is presented in three parts. It is concerned with, respectively, the process of interpretation, illustrative analyses, and description of the pictures. There is an appendix which includes six pages of notes on the administration of the test and a very comprehensive and specifically classified bibliography of the thematic apperception technique, which takes up nearly 30 pages. Though well written, it does assume interest and persistence on the part of the reader as well as some familiarity with the background and use of the thematic apperception tests for the study of personality as a whole as well as fantasy. It is a fact, as the author says, that the interpretive significance of the themas perceived and stories told are derived from the understanding of the ways in which the emotional preoccupations are symbolized in words and plots, from the ways in which emotions bow to propriety in overt expression, or ways in which each individual lessens his own burden by expressing his own feelings through disguising them for himself and others. To the area of personality where fantasy rules, logic and propriety do not apply, and it is through the analysis of a manner in which the person manipulates the symbols of conventional communication that one derives insight into personal motives. How one learns or gains knowledge about skills directly and feelings indirectly in the use of the test is considered at some length.

Variables of form and content are treated in separate chapters.

Content refers to what the subject said and form to how he said it. And both categories are set in a conceptual framework for individual case analysis wherein comparisons are made with Rorschach formulations. The cases presented include a record of a healthy, middle class boy, a 29-year-old man, a Dakota Sioux of 14, and a Navaho girl of 11. In the chapter on the stimuli properties of the pictures, each card is described under six captions: Murray's Description; Manifest Stimulus Demand; Form Demand; Latent Stimulus Demand; Frequent Plots and Significant Variations.

This is a highly worthwhile volume for all people interested in the use of projective technique in the study of personality and for the thematic apperception tests so used particularly. Investigators of any phase of problems related to the validation or the use of the thematic apperception technique will find the bibliography very helpful.

H. Meltzer
Human Relations Research Foundation, St. Louis, Missouri

Otto Pollak. *Integrating Sociological and Psychoanalytic Concepts: An Exploration in Child Psychotherapy*, New York: Russell Sage Foundation, 1956. Pp. 284.

Effective application of social science concepts and approaches can make a significant contribution towards the improvement of diagnosis and therapy in child guidance work. This is the author's thesis, and the material used and the interpretations made are all slanted in the direction of proving this thesis. To a fair extent, the whole story is the report of Dr. Pollak's experience with a child guidance team. A clinical psychologist, because of illness, was not included in the team.

In this volume, the author raises a number of interesting problems that are relevant in considering ways of improving child guidance efforts and results. The much to-do about the neglect of the father in child guidance work is not new but is relevant. The concepts he advocates the use of, which in the author's way of thinking represents a reformulation of method, are four in number: family orientation, social interaction, social role, and culture conflict.

The content of the volume is presented in four parts. Part I, which is called "Family Pathology, Family Diagnosis and Family

Therapy with Child Welfare Focus," is illustrated with the presentation of two case studies. Part II, called "The Therapeutic Management of Extrafamilial Factors" is illustrated with three case studies. Part III, called "A Research Outlook" covers psychiatric literature and differences between theory and practice. Part IV is called "The Agency Looks at the Project" and is an evaluative comment by Herschel Alt, Executive Director of Jewish Board of Guardians, which, with the Russell Sage Foundation, made possible this project.

Some of the concepts stressed by Pollak are worthy of consideration. Some are over-generalizations. For example, when he says that a child welfare focus would result in consideration of situations rather than individuals so that we would not have sick husbands or sick wives, but sick marriages; not sick children or sick mothers or sick fathers, but sick families; not sick teachers or sick students, but sick school situations; not immature, sick guardians or sick administrators, but sick institutions, he is calling attention to the advisability of emphasizing emotional climate of groups rather than emotional climate of individuals; but how much reality is he running away from, as well as getting into, by so doing? It may be a fact that the "bad child" so-called, referred to by the mother as such, is best conceived as being in part made that way by the mother and all things in the form of human surroundings at home. It doesn't follow, however, that in all instances there are no such things as sick individuals if some institutions do permit, almost invite, sickness in the behavior of the people who live in them. It also doesn't follow that even in such instances that some people who come there have been made sick before they came there. The universal implication of the author's emphasis would be to have an institution to house institutions rather than individuals. To use a word which appeared in the "Talk of the Town" in a recent issue of the *New Yorker*, the author has done too much "jumboizing" of his concepts and by so doing has run off the track of reality, as well as made vivid some realities deserving attention.

On the assumption that things well done should be given survival value rather than be reformulated or thrown out, there are surprising omissions here. Included are: No mention of the work of Healy or Fritz Redl; no attention to the organization of abilities and attitudes towards the use of them; no consideration

of economic factors. A greater awareness of the work realities of the clinical team and a lesser emphasis on verbal reformulations would more likely yield a more relevant and enriched exploration in child psychotherapy.

H. MELTZER
Human Relation Research Foundation, St. Louis, Missouri

FRANK A. LOGAN, DAVID L. OLMSTED, BURTON R. ROSNER, RICHARD D. SCHWARTZ, & CARL M. STEVENS. *Behavior Theory and Social Science*, New Haven: Yale University Press, 1955.

A group of Ford fellows in behavior science with a three-year appointment at the Institute of Human Relations of Yale University have produced this volume which is an attempt to restate the Hullian point of view with regard to behavior theory and to make applications of this theory to problems in social science. The five authors of this volume, according to their own statement, met as a group five days a week throughout the first year for several hours a day and less frequently during the second and third years in which the report was written. They also acknowledge their indebtedness to a large number of other colleagues at Yale with whom they could discuss Hull's point of view and its applications.

It is indeed significant that this is one of a number of books to appear in the past two years to present an extended analysis of Hull's point of view. There has been a kind of race to see which of the theories propounded by Guthrie, Hull, Skinner, and Tolman would have the pre-eminence. It is not certain by any means that the issue has been finally settled and history will probably record that each of these psychologists has made his enduring contribution. Certainly at the present time, however, Hull's theoretical structure intrigues the largest number of psychologists and there has been greater elaboration of his theories by other psychologists than of any of the other current learning and behavior theories.

One chapter in the book is devoted to discussing the properties of the response, and another chapter is given to discussion of the stimulus and other antecedent conditions. In chapter 5, entitled "Language," the writers go beyond the basic principles under-

lying Hull's point of view to a discussion of thinking and problem solving. Chapter 6, entitled "Free Behavior," deals with the same phenomena that Skinner discusses under the title of operant behavior; that is, behavior which is in response to stimuli not immediately present. These authors prefer to call this free behavior, and they make a strenuous effort to fit it into the Hullian framework. Since the results are somewhat speculative, they have drawn up three models of such behavior and then state that these should be the subject for much experimentation in order to determine which of the three models comes closest to fitting actual experience. The last two chapters in the book deal with social interaction and cultural behavior. Here there is still more speculation, but always attempting to derive an explanation of social behavior from the basic Hullian proposition.

This reviewer can recommend this book to serious students of learning and behavior theory as a discussion which moves our theoretical formulation a step ahead.

PERCIVAL M. SYMONDS

Teachers College, Columbia University

PUBLICATIONS RECEIVED

GUY L. BOND AND MILES A. TINKER. *Reading Difficulties: Their Diagnosis and Correction.* New York: Appleton-Century-Crofts, Inc., 1957, pp. 486. $5.25.

WILLIAM F. BRUCE AND A. JOHN HOLDEN, JR. *The Teacher's Personal Development* New York: Henry Holt & Co., 1957, pp. 346. $4.25.

MARGARET DOREN. *Doren Diagnostic Reading Test of Word Recognition Skills.* Minneapolis: Educational Publishers, Inc., 1956, pp. 24 (paper).

ELAINE CUMMING AND JOHN CUMMING. *Closed Ranks—An Experiment in Mental Health Education.* Cambridge: Harvard University Press, 1957, pp. 192. $3.50.

LUCIEN DINTZER. *Le Jeu D'Adolescence.* Paris: Presses Universitaires de France, 1956, pp. 123 (paper).

WALTER CROSBY EELLS, Editor. *College Teachers and College Teaching: An Annotated Bibliography on College and University Faculty Members and Instructional Methods.* Atlanta, Georgia: Southern Regional Education Board, 1957, pp. 282 (paper).

JOHN EISENSON AND MARDEL OGILVIE. *Speech Correction in the Schools.* New York: The Macmillan Co., 1957, pp. 294. $4.25.

CLAUDE FRANÇOIS-UNGER. *L'Adolescent Inadapté: Réadaptation Sociale et Formation Professionnelle.* Paris: Presses Universitaires de France, 1957, pp. 302 (paper).

SOL L. GARFIELD. *Introductory Clinical Psychology.* New York: The Macmillan Co., 1957, pp. 469. $6.00.

DAVID GREENWOOD. *Truth and Meaning.* New York: Philosophical Library, 1957, pp. 114. $3.75.

STEN HENRYSSON. *Applicability of Factor Analysis in the Behavioral Sciences: A Methodolical Study.* Stockholm: Almquist & Wiksell, 1957, pp. 156 (paper).

DALE B. HARRIS, Editor. *The Concept of Development: An Issue in the Study of Human Behavior.* Minneapolis: University of Minnesota Press, 1957, pp. 287. $4.75.

ROBERT L. KAHN AND CHARLES F. CANNELL. *The Dynamics of Interviewing.* New York: John Wiley & Sons, 1957, pp. 368. $7.75.

FREDERICK KILANDER. *Health for Modern Living.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1957, pp. 494. $4.95.

GRACE LANGDON AND IRVING W. STOUT. *Helping Parents Understand Their Child's School: A Handbook for Teachers.* Englewood Cliffs, New Jersey: Prentice-Hall, 1957, pp. 508.

EDNA S. LEVINE. *Youth in a Soundless World: A Search for Personality.* New York: New York University Press, 1956, pp. 217. $5.00.

MARIAN EAST MADIGAN. *Psychology: Principles and Applications.* Second Edition. St. Louis: C. V. Mosby Co., 1957, pp. 391. $4.75.

D. J. O'CONNOR. *An Introduction to the Philosophy of Education.* New York: Philosophical Library, 1957, pp. 148. $3.75.

C. W. ODELL. *A First Course in Educational Statistics.* Dubuque, Iowa: Wm. C. Brown Co., 1957, pp. 172 (paper). $3.75.

HAROLD PALMER. *Psychopathic Personalities.* New York: Philosophical Library, 1957, pp. 179. $4.75.

E. LAKIN PHILLIPS AND JAMES F. GIBSON. *Psychology and Personality.* Englewood Cliffs, New Jersey: Prentice-Hall, 1957, pp. 338. $4.95.

SIDNEY L. PRESSY AND RAYMOND G. KUHLEN. *Psychological Development Through the Life Span.* New York: Harper & Brothers, 1957, pp. 654. $6.00.

Regional Problems and Issues in Human Relations Education. New York: Anti-Defamation League of B'nai B'rith, pp. 78 (paper).

HERBERT A. SIMON. *Models of Man: Mathematical Essays on Rational Behavior in a Social Setting.* New York: John Wiley & Sons, Inc., 1957, pp. 287. $5.00.

D. H. STOTT. *Unsettled Children and Their Families.* New York: Philosophical Library, 1957, pp. 240. $6.00.

W. L. SUMNER. *Visual Methods in Education.* Second Edition. New York: Philosophical Library, 1957, pp. 231. $6.00.

PERCIVAL M. SYMONDS. *Dynamics of Psychotherapy: The Psychology of Personality Change. Volume II. Process.* New York: Grune & Stratton, 1957, pp. 398, $6.50.

MILDRED C. TEMPLIN. *Certain Language Skills in Children: Their Development and Interrelationships.* Minneapolis: University of Minnesota Press, 1957, pp. 183. $4.00.

ARTHUR E. TRAXLER. *Techniques of Guidance.* Revised Edition. New York: Harper & Brothers, 1957, pp. 374. $6.00.

CLARK E. VINCENT. *Readings in Marriage Counseling.* New York: Thomas Y. Crowell Co., 1957, pp. 500. $4.50.

CHARLES H. WILSON. *A Teacher Is a Person.* New York: Henry Holt and Co., 1956, pp. 285. $3.75.

ARTHUR WORMHOUDT. *Hamlet's Mouse Trap: A Psychoanalytical Study of the Drama.* New York: Philosophical Library, pp. 221. $3.50.